I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17976

Measuring the Sources of Taste-Based Discrimination Using List Experiments

Ariel Listo
Ercio A. Muñoz
Dario Sansone

JUNE 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17976

Measuring the Sources of Taste-Based Discrimination Using List Experiments

Ariel Listo
University of Maryland

Ercio A. Muñoz
Inter-American Development Bank

Dario Sansone
University of Exeter and IZA

JUNE 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Measuring the Sources of Taste-Based Discrimination Using List Experiments*

This paper examines how attitudes among supervisors, co-workers, and customers are related to discrimination against sexual minority individuals in the workplace. Participants from a large, nationally representative online sample in Chile took part in double list experiments – which reduce social desirability bias when eliciting views on sensitive topics – followed by direct questions on attitudes toward sexual minority individuals. The findings reveal a discrepancy between reported and actual levels of comfort with gay individuals in the labor market. The respondents underreported their discomfort by 15-23 percentage points, with the largest bias and lowest comfort levels observed when they were asked about supervising gay employees. These attitudinal patterns were mirrored in incentivized donation behavior: individuals who chose not to donate any amount from a lottery to a local LGBTQ-related nonprofit reported lower comfort levels and exhibited greater misreporting. Finally, the respondents consistently underestimated the broader societal support for gay employees and co-workers.

JEL Classification: C93, D91, J15, J71, Z13

Keywords: LGBTQ+, Chile, discrimination

Corresponding author:
Dario Sansone
University of Exeter
Stocker Rd
Exeter EX4 4PY
United Kingdom
E-mail: d.sansone@exeter.ac.uk

* Financial support from the Inter-American Development Bank (ESW RG-E1952) is gratefully acknowledged. The views expressed in this paper are those of the authors and should not be attributed to the Inter-American Development Bank. We would like to thank Caridad Araujo, Anne Boring, Keith Marzilli Ericson, Adrian Rubli, and Javiera Selman for their helpful comments. We also thank the seminar participants at the AEA LGBTQ+ Virtual Seminar Series, the Bristol Applied Economics Meetings, the European Commission (Joint Research Centre), the University of Bath, the XXI Ridge Forum LACEA Brain, and the University of Oxford for their feedback. All errors are our own.

1. Introduction

Understanding the sources and drivers of discrimination carries important policy implications for reducing inequality and improving the allocation of human capital. For instance, for statistical discrimination due to incorrect beliefs, informational campaigns may be necessary to correct such beliefs. In contrast, taste-based discrimination that is rooted in employers' preferences calls for appropriately enforced anti-discrimination laws, whereas such laws may be less effective against customer-driven taste-based discrimination. Therefore, to effectively inform policymakers, researchers need to not only uncover evidence of discrimination against different groups but also analyze its causes.

This paper examines how attitudes among supervisors, co-workers, and customers contribute to workplace taste-based discrimination against sexual minority individuals by drawing on a large, nationally representative online sample in Chile. Specifically, it addresses the following questions: What is the level of comfort with sexual minority individuals in the workplace among supervisors, co-workers, and customers? Is there evidence of misreporting in these self-reported attitudes? Do individuals underestimate the broader societal support for sexual minority individuals? Finally, do these stated attitudes influence real-world behaviors?

Measuring attitudes toward minority individuals is itself policy relevant: attitudes can affect health and socioeconomic behaviors, outcomes, and disparities (Aksoy, Chadd, and Koh 2023; Glasman and Albarracín 2006; NASEM 2020), as well as occupational sorting (Plug, Webbink, and Martin 2014; Gutiérrez and Rubli 2024b). Attitudes can also directly induce minority stress, that is, stress due to internalized homophobia and transphobia, anticipated rejection, constant efforts to hide one's identity, and actual experiences of discrimination and violence (Meyer 1995). In addition, while there is evidence of positive effects of employment anti-discrimination laws (Donohue and Heckman 1991; Klawitter and Flatt 1998; Neumark and Stock 2006; Klawitter 2011; Button 2018; Neumark et al. 2019), the effectiveness of such employment protections depends on compliance and the level of support that they receive: if employers have a high distaste for sexual minority individuals (or if they believe that other employees or customers may dislike interacting with such individuals), they will try to find ways to circumvent these laws. Relatedly, support for certain groups or policies may actually impact voting behavior (Friese et al. 2012; Castanho Silva, Fuks, and Tamaki 2022).[1]

One key methodological challenge in measuring attitudes toward sexual minority individuals in the workplace is that, when asked directly, people may misreport their preferences and beliefs. For instance, misreporting may be driven by participants fearing that their truthful answers will not be socially acceptable. This phenomenon is usually referred to as social desirability bias or sensitivity bias (Blair, Coppock, and Moor 2020). These attitudes are also hard to detect from actual behavior since employers', co-workers', and customers' choices to hire or interact with sexual minority individuals may be influenced by a wide range of factors,

[1] For instance, (Stephens-Davidowitz 2014) argued that indirect measures of local racial animosity via Google Trends were larger than estimates obtained from direct survey questions and correlated with voting results in US presidential elections.

such as the number and quality of available alternatives, legal constraints, and beliefs about differential productivity between sexual minority and heterosexual individuals.

This paper overcomes these barriers by relying on a list experiment: individuals are asked to report how many of the statements presented in a list are true for them (without sharing which specific statements are true for them). The control group of respondents is presented with a short list of non-key statements, while the treatment group is presented with the same short list plus an additional key sensitive statement. In this survey, the key sensitive statements pertain to the participants' level of comfort supervising a gay employee, working closely with a gay co-worker, or having a cashier at the supermarket who is gay. Comparisons of the average number of items reported across lists make it possible to estimate the true share of respondents who agreed with each key statement of interest regarding sexual minority individuals. More specifically, as detailed in Section 2.1, this study relies on double list experiments (i.e., two parallel list experiments) to verify the robustness of the findings to the use of different non-key statements (Chuang et al. 2021) and to increase the precision of the estimates (Droitcour et al. 1991; Glynn 2013).

While the list experiment technique cannot identify which specific individuals agree with the key statements (because individuals report only the total number of statements within each list that are true for them, as opposed to indicating whether each statement is true for them), it allows researchers to credibly compute population-level estimates regarding sensitive issues while reducing social desirability bias. It is then possible to estimate the magnitude of such social desirability bias through comparisons of the estimates from the list experiments to the average responses to questions directly asking individuals about key sensitive issues.

The main results of the double list experiments indicate that individuals underreported their discomfort with gay individuals in the labor market by 15 to 23 percentage points. The highest gap between the reported comfort levels and the list experiment estimates arises in the supervisor scenario, in which the participants reported the lowest level of comfort with gay employees. However, high misreporting rates are also detected when measuring levels of comfort with a gay co-worker or a gay cashier. Notably, the respondents were specifically asked about their level of comfort – not whether they thought a gay employee, co-worker, or cashier would be more productive – to focus the analysis on taste-based discrimination rather than on statistical discrimination.

To link these results to actual, real-stakes behavior, the respondents were entered into a lottery with a chance to win approximately USD 100 extra (roughly 20% of the minimum monthly wage). Before the results of the lottery were announced, the participants were asked, if they were to win the lottery, how they would allocate the extra funds between themselves and a local nonprofit organization promoting LGBTQ+ equality. The estimated differences in discomfort from the list experiments and direct questions correlate with real-stakes donation behavior: Among individuals who chose not to donate any amount to an LGBTQ-related nonprofit organization, the degree of misreporting is higher, and the level of comfort is lower.

The survey experiments were supplemented by standard sociodemographic questions and additional opinion questions. These questions allow for comparisons between the respondents' level of support for gay individuals and their levels of support for other sexual minority individuals, as well as their levels of support for other minority individuals and in different environments. The findings from these questions suggest that the stated level of comfort tends to be lower for interactions with gay men in the workplace compared to other contexts or other minority individuals. In addition, the empirical section of this paper includes heterogeneity analyses that estimate the level of comfort and social desirability bias among specific subgroups based on observed characteristics such as sex, education, income, employment status, managerial experience, political affiliation, or religiosity. These estimates indicate higher levels of comfort among women, those with managerial experience, those with left-leaning political affiliations, and non-religious respondents. However, misreporting is widespread across most groups.

In addition, the respondents were asked to estimate the shares of the Chilean adult population who would be comfortable with a gay employee, co-worker, or cashier. These shares can be compared to the estimates from the list experiments to test whether individuals underestimated or overestimated the level of comfort with sexual minority individuals in the general population. Individuals' beliefs about population attitudes toward gay employees and co-workers are consistently lower than their own elicited support, even after accounting for social desirability bias.

These analyses contribute to the general literature on labor market discrimination, to the more specific literature on economic disparities by sexual orientation and gender identity, and to the interdisciplinary literature that employs list experiments. A large share of studies on the economics of gender, race, disability, sexual orientation, and gender identity, as well as other legally protected characteristics, focus on testing whether there is any evidence of discrimination in the labor market. These studies predominantly rely on observational data (Blau and Kahn 2017) – via methodologies such as the Kitagawa–Blinder–Oaxaca decomposition – or on correspondence experiments (Neumark 2018; Lippens, Vermeiren, and Baert 2023; Kessler, Low, and Shan 2025) – that is, sending fictitious comparable CVs with varying selected features to real employers with job openings. A much smaller number of studies try to link patterns of discrimination to a specific theory (Bertrand and Duflo 2017; Bohren et al. 2025): almost all of these studies aim to compare statistical discrimination (Arrow 1973) and taste-based discrimination (Becker 1971), although other studies have expanded these models or examined different kinds of discrimination such as institutional or systemic discrimination (Small and Pager 2020; Bohren, Hull, and Imas 2025; Onuchic 2024).

An even more limited set of studies attempts to uncover the potential source of taste-based discrimination (Dolado, Minale, and Guerra 2023), that is, whether the discrimination stems from employers', co-workers', or customers' preferences not to interact with members of a particular group.[2] This study relies on a series of double list experiments to advance this

[2] For instance, (Borm et al. 2020) found suggestive evidence from a lab experiment with student participants of co-worker and customer taste-based discrimination against transgender workers but not of employer taste-based

literature focused on identifying and measuring the sources of taste-based discrimination in the context of discrimination against sexual minority individuals in the workplace in Chile. Importantly, it exploits a large, weighted representative sample, one of the largest ever used to conduct list experiments.

While there is a substantial literature analyzing discrimination based on gender, age, race, and disability (Goldin 2014; Blau and Kahn 2017; Neumark 2018; Goldin 2021; Lippens, Vermeiren, and Baert 2023), the number of studies focusing on LGBTQ+ discrimination is significantly smaller, although rapidly rising (Sansone 2019; Badgett, Carpenter, and Sansone 2021; Badgett et al. 2024; 2025). This growth in LGBTQ+ data and research aligns with the increasing share of individuals identifying as LGBTQ+ (Jackson 2023). Most previous studies have found large inequalities in the workplace affecting sexual and gender minority people: for example, wage penalties for gay men and bisexual individuals have been documented in numerous countries, and LGBTQ+ individuals are less likely to be invited for job interviews (Badgett et al. 2024). This study advances this literature by investigating potential drivers of these documented labor market disparities.

List experiments have been used extensively in other social sciences, such as sociology and political science (Rayburn, Earleywine, and Davison 2003; Lax, Phillips, and Stollwerk 2016; Gervais and Najle 2018; Streb et al. 2008; Blair, Coppock, and Moor 2020; Li and Van den Noortgate 2022), including in Chile to study vote buying (de Jonge 2015). However, they have been used less frequently in economics (Coffman, Coffman, and Ericson 2017; Aksoy, Carpenter, and Sansone 2025; Jamison, Karlan, and Raffler 2013; Chuang et al. 2021; Boring and Delfgaauw 2024; McKenzie and Siegel 2013; Agüero and Frisancho 2022).[3] List experiments are particularly effective in measuring attitudes – as in this study – rather than behavioral or personal characteristics (Ehler, Wolter, and Junkermann 2021). In addition, as already emphasized by (Osman, Speer, and Weaver 2025), the use of list experiments to study discrimination in countries outside Western Europe and North America with limited administrative data can be particularly valuable since observational analyses via Kitagawa–Oaxaca–Blinder decomposition are restricted by what variables can be directly observed. Furthermore, correspondence experiments are less feasible in economies with many small firms that do not rely on resumes or job websites, thus underestimating the extent of discrimination, as small firms are the enterprises most likely to be biased.

Chile is an ideal context for studying these issues for several reasons. First, with a few exceptions (Muñoz and Sansone 2024; Muñoz, Saavedra, and Sansone 2024; Muñoz, Sansone,

discrimination. Relatedly, (Bar and Zussman 2017) showed that Jewish Israelis are willing to pay a premium to receive services from Jewish rather than Arab workers, while (Kelley et al. 2025) provided evidence of gender-based customer discrimination in a randomized field experiment that varied the names of online sales agents.

[3] There are two recent list experiments worth highlighting since they were also conducted in Latin America to estimate the size of the LGBTQ+ population: (Ham, Guarin, and Ruiz 2024) in Bogotá, Colombia, and (Gutiérrez and Rubli 2024a) in Mexico. (Gutiérrez and Rubli 2024a) used a list experiment to also test whether people would rather work with a straight person, but their estimate from the list experiment was counter-intuitive and lower than the average from the direct question: the authors acknowledged that these findings may have been driven by people misinterpreting the question. Extensive piloting and cognitive interviews were conducted to avoid encountering the same problem.

and Ysique 2024; Nettuno 2024; Tampellini 2024; Muñoz, Saavedra, and Sansone 2025), most LGBTQ+ studies have focused on high-income countries.[4] Therefore, there is an urgent need to reduce the historical invisibility of LGBTQ+ individuals in the Global South. Second, Chile has a developed formal labor market, which facilitates the study of workplace discrimination (e.g., co-worker preferences). Third, Chile has made significant legislative advances in LGBTQ+ rights in recent decades, especially after the end of the military dictatorship in 1990: same-sex sexual activities were decriminalized in 1999, an employment anti-discrimination law was passed in 2012, and same-sex marriage was legalized in 2021. At the same time, attitudes toward LGBTQ+ individuals are also generally positive: most people support the right of same-sex couples to get married and think that sexual minority individuals should be permitted to run for public office (Muñoz, Sansone, and Ysique 2024). Similarly, 78 percent of the respondents in this study report a gay acquaintance in their social network (Table B1). However, there is evidence of recent backlashes (Palacios 2024), while gender norms remain conservative, and gender-based discrimination is still widespread (Montoya et al. 2025). Thus, it remains important to understand whether LGBTQ+ support is context specific or is limited to only certain sexual and gender minority individuals.

2. Data and methodology

2.1. List experiments and survey design

The main analysis relied on the list experiment technique (also called the "item-count technique," "unmatched count," or "veiled approach"), which was pioneered by (Raghavarao and Federer 1979) and (Miller 1984). In this technique, respondents are given a list of statements and are asked to report how many (but not which specific) statements are true for them, thus providing an extra layer of anonymity and increasing privacy (Coutts and Jann 2011). In this study, the participants were assigned to either a treatment group or a control group. In the control group ("short list"), the participants were given a list of statements and were asked to indicate how many of those statements were true for them. In the treatment group ("long list"), the participants were given the same list of statements plus a key statement of interest (in this study, a statement about their level of comfort with gay individuals in different contexts). The difference in means between the two lists can be interpreted as the estimated share of the population with the key attribute of interest. Table 1 presents one of the lists used in this study (translated from the original Spanish version).

This technique can then be extended by employing double list experiments. In this study, for each key sensitive statement, the participants were presented with two lists (List A and List B) whose items were designed to be positively correlated. Each list contained four non-key statements. Half of the participants (randomly selected) saw List A (a short list) and then List B with the key statement (a long list). The other half saw List A with the key statement (a long list) and List B (a short list). The differences in means between the short and long lists from both Lists A and B were averaged to provide the estimated share of the population with that

[4] Another study worth mentioning is (Abbate et al. 2024): these authors conducted a correspondence experiment in the rental housing market in four Latin American countries and found evidence of statistical discrimination against couples with a transgender woman.

key sensitive attribute. Owing to this extension, it was possible to obtain more precise estimates by increasing power and reducing variance since all respondents provided information about all key statements (akin to the advantages of a within-subject design), unlike a single list experiment in which only respondents who see the long list provide such information (Droitcour et al. 1991; Glynn 2013). An additional advantage of the double list method is that it made it possible to verify the robustness of the main findings to the use of different non-key statements by comparing the estimates obtained from List A against those obtained from List B (Chuang et al. 2021).

This study included three sets of double list experiments with three key statements:

Supervisor preference: "I would feel comfortable supervising a gay employee."

Co-worker preference: "I would feel comfortable working closely with a gay co-worker."

Customer preference: "I would feel comfortable having a cashier at the supermarket who is gay."

The double list experiment technique was employed for all three statements, thus leading to a total of six lists: Lists 1A and 1B for supervisor preference, Lists 2A and 2B for co-worker preference, and Lists 3A and 3B for customer preference.[5] Importantly, because Spanish is a gendered language, these statements implicitly treat gay men, rather than lesbian women, as the reference category. Comparisons with attitudes toward lesbian women are discussed in Section 3.4.

Direct questions regarding the key statements were then asked to all participants after they responded to demographic and socioeconomic questions in a questionnaire. The direct questions provided baseline estimates of the share of the population with the key attributes, thus making it possible to estimate the size of social desirability bias. Ex ante, the size of this bias is not clear: online surveys may elicit truthful answers since they are self-administered, potentially completed in private, and anonymous (Holbrook and Krosnick 2010; Robertson et al. 2018). Thus, the magnitude of misreporting documented in this study is likely to be a lower bound to what might occur in other surveys since most surveys are not conducted with as much privacy and anonymity and, thus, the respondents in this study may have been less prone to social desirability bias even when answering direct questions.

All participants first completed the list experiment section and then advanced to the survey. They were not allowed to skip any questions in the list experiments and were not allowed to go back and revise their answers at any point. In addition to the three direct questions (related to the three key statements from the list experiments) and standard demographic and socioeconomic variables, the questionnaire included items regarding sex, sexual orientation, gender identity, and additional direct questions to measure the participants' stated views on women and minority individuals in different environments.

[5] Based on the 2024 Chilean Census, the share of male workers in the services industry, including sellers in commerce and markets (e.g., cashiers) was approximately 42 percent.

The participants were then asked to provide their estimates of the share of the Chilean adult population who would feel comfortable supervising a gay employee, working closely with a gay co-worker, or having a cashier at the supermarket who is gay. Finally, the participants were entered into a lottery and were asked to indicate how much they would like to donate from their potential earnings to a local LGBTQ+ nonprofit if they were among the randomly selected winners (each receiving CLP 100,000, i.e., approximately USD 100).[6]

2.2. Data collection and study sample

The first draft of the questionnaire, list experiments, and pre-analysis plan were reviewed by academic experts between June and July 2024. The updated experiment protocol was pre-registered in July 2024, followed by ethical approval from the authors' universities.[7]

Subsequently, a pilot study was conducted on an online platform (Prolific) via a non-representative sample of 535 respondents from Chile. The findings from this pilot study are discussed in Section C of the Online Appendix. Notably, the pilot study provided evidence that completing the list experiments first did not affect the answers to the direct questions on attitudes toward sexual minority individuals (Appendix C.1), it provided guidance on the appropriate terminology to use in Chile (Appendix C.2), and it confirmed that the survey instructions were clear and that the participants were not primed to think about LGBTQ+ issues (Appendix C.3).

In addition, the answers in the pilot study supported the claim that the expressed levels of comfort with gay men in the contexts reviewed in this study were driven by the participants' own distaste or preferences (i.e., taste based) rather than by productivity beliefs (i.e., statistical), as intended in the survey design (Appendix C.4). This result is important and reassuring, as the specific wording ("comfortable") and occupation chosen to elicit customer preferences ("cashier") were selected specifically to minimize the risk of productivity-related expectations or beliefs influencing the participants' answers. The choice of cashier was also influenced by findings from other studies noting customers' distaste for LGBTQ+ individuals handling their groceries (Webb 2025).

Following this pilot study, a local survey company in Chile, Datavoz,[8] conducted 10 cognitive interviews in August 2024. Datavoz then conducted a pilot study with 235 participants, followed by a soft launch with 62 participants. The main data were collected by Datavoz from

[6] Section A.1 of the Online Appendix discusses ethics and pre-registration. Section A.2 presents additional technical details on the list experiment technique and the sources of sensitivity bias. Section A.3 discusses the randomization of the items within each list experiment, as well as the randomization of the order of the list experiments: there is little evidence that the answers in the list experiments were affected by the order in which the participants saw the lists. Section A.4 discusses several list experiment design considerations (e.g., selecting the number and type of non-key items and avoiding priming the respondents), provides evidence that floor and ceiling effects were negligible in these list experiments, and describes additional advantages of survey experiments and online surveys. Section A.5 provides additional evidence supporting the validity of the list experiment assumptions (treatment randomization, no design effect, and no liar).

[7] See https://www.socialscienceregistry.org/trials/13989.

[8] Datavoz is an established survey company that is routinely used by organizations such as the United Nations Development Programme, the International Labour Organization, Vanderbilt University, and the Universidad Católica de Chile.

early October until mid-December 2024, and the survey included 4,000 participants. The final sample excludes data from the pilot study, soft launch, and cognitive interviews and includes data from only the main wave.[9]

Importantly, the final sample size is larger than that in most previous studies. In fact, almost all the list experiments summarized by (Blair, Coppock, and Moor 2020) had fewer than 1,000 respondents. Similarly, only 4 out of the 54 list experiments reviewed by (Li and Van den Noortgate 2022) had more than 1,000 observations in the control group (short list) and 1,000 in the treatment group (long list). This sample feature is likely to have been the key factor in some of these studies not having enough power to detect differences in prevalence rates or social desirability bias, especially across subgroups.[10]

The participants in the main study were recruited based on sex, region, and age targets with the goal of obtaining a weighted sample that was representative of the Chilean population along these three dimensions. The questionnaire and list experiments were coded on Qualtrics. The survey was tested to work effectively and easily on multiple platforms (Windows, Apple, and Android), and it was designed to be mobile friendly. The recruitment email and survey used by Datavoz are presented in Section E of the Online Appendix. The participants never disclosed any identifying information, and the survey was completely anonymous. The main empirical analyses include all respondents who finished the experiment and the survey in their entirety.[11]

The main questionnaire took approximately 14 minutes to complete (median response time). Fifty gift cards in the amount of around CLP 50,000 (USD 50) each were then raffled off among the participants who completed the survey as participation payments. Additionally, up to 50 additional gift cards in the amount of roughly CLP 100,000 (USD 100) were raffled off among all respondents as part of the donation question.

All variables used in the main empirical analysis are described in Appendix B. In addition, Table B1 presents the weighted summary statistics for the Datavoz participants. A comparison of the sample to official population estimates from the census and other national representative surveys shows that the main weighted sample is representative not only based on age, region, education, and sex – as expected, given the sampling and weighting methodology (discussed in Appendix A.6) – but also broadly with respect to several other variables such as religious affiliation (32 percent of respondents reported no religious affiliation, compared to 29 percent of respondents from 2023 Latinobarometro from Chile), political affiliation (55 percent of respondents reported that they "lean left," compared to 47 percent of respondents from the 2018 World Value Survey – Chile), and indigenous status (10 percent of respondents reported

[9] Unweighted results, including data from the pilot study and the soft launch, are included in Section A.6 of the Online Appendix. The main findings remain qualitatively similar, with a slightly larger level of support estimated by the double list experiment with customers, resulting in a slightly lower estimated social desirability bias for that group.

[10] One exception is the study in Mexico by (Gutiérrez and Rubli 2024a), which had a larger sample size (approximately 10,000), given their focus on measuring LGBTQ+ identification.

[11] Section A.6 in the Online Appendix reports additional data quality checks, such as checking that the participants paid attention, excluding respondents who provided the same answer in all list experiments, excluding respondents who completed the survey too quickly or too slowly, excluding participants who scored poorly in the Captcha Verification test, or using unweighted data. This section also discusses how standard errors for the main analysis were computed.

indigenous descent, compared to 13 percent of respondents from the 2017 Chilean Census). Similarly, the weighted average income range is between CLP 1,200,001 and CLP 1,450,000 (the median income range is between CLP 975,000 and CLP 1,2000,000); thus, it is close to the average income of CLP 1,304,771 estimated in the 2022 Encuesta de Caracterización Socioeconómica Nacional (CASEN). Moreover, the estimated share of sexual and gender minority individuals is consistent with that in previous online surveys (Jackson 2023). Regarding employment, 80.6 percent of respondents in the weighted sample were considered employed, while the 2018 World Value Survey found that 78.3 percent were employed or retired/pensioners.

In addition, as reported in the next section, this study finds that even in the list experiments, the majority of respondents exhibit support for gay people. This finding is in line with the results of most surveys conducted in Chile over the last decade. Approximately 60 percent of Latinobarometro's 2023 respondents answered that they "agree" or "strongly agree" with the following statement: "Homosexual couples are just as good parents as other couples." This support has been increasing over time: in response to the same question in the 2018 World Value Survey, 35.9 percent of respondents answered that they "agree" or "strongly agree." In the same 2018 wave, 72.3 percent of participants did not mention "homosexuals" out of a list of groups of people whom they would not like to have as neighbors. Similarly, 67.4 percent of respondents "strongly approved" or "approved" of same-sex couples having the right to marry according to the 2023 AmericasBarometer.

One may argue that the main drawbacks of these kinds of online samples are that it is difficult to estimate the margin of error for the general population and that they do not include respondents from among the population of internet users. However, as noted by (Haaland, Roth, and Wohlfart 2023), given that most probability-based panels have relatively high nonresponse rates, the differences in the extent of selection between probability-based samples and quota-based online datasets might not be very large in practice. Moreover, (Haaland, Roth, and Wohlfart 2023) summarized evidence from several studies arguing that the online and offline populations hardly differ in terms of survey responses and experimental results. It is also reassuring to note that 90 percent of the Chilean population used an internet connection in 2021, which is comparable to 92 percent in the US (World Bank 2022).

3. Results

This section consists of five parts. First, it presents the main results from the double list experiments. Second, it discusses how the estimates from the list experiments compare with population beliefs. Third, it conducts heterogeneity analyses based on a set of relevant participant characteristics. Fourth, it compares the results on stated attitudes toward gay individuals to attitudes toward other sexual minority individuals and other minority groups. Finally, it maps the findings from the double list experiments and from the direct questions onto behavior via a real-stakes, incentive-compatible donation question.

3.1 Main results from the list experiments

The main results stem from a comparison of the findings from the double list experiment and from the direct questions on attitudes among supervisors, co-workers, and customers. These

weighted results are presented in Figure 1. The first two bars (from left to right) of Figure 1 measure attitudes toward supervising a gay employee based on the answers to the list experiments (dark gray bar labeled "Double List") and the direct question (light gray labeled "Direct Question"). The middle two bars calculate attitudes toward a gay co-worker. The last two bars estimate attitudes toward a gay cashier. In each of these three sets of dark and light gray bars, the difference between the estimates from the list experiments and from the direct question is reported above the bars. For each category, the double list experiment share was obtained by averaging the estimates from two list experiments with different non-key items (List A and List B). The weighted and unweighted estimates from the single list experiments are separately reported in Table D1.

All three sets of bars indicate that supervisors', co-workers', and customers' discomfort with gay individuals is substantially underreported by approximately 15 to 23 percentage points. The first two bars in Figure 1 indicate that the estimated share of individuals who reported that they are comfortable supervising a gay employee is 57.3 percent, whereas the share of individuals who reported that they are comfortable supervising a gay employee in the direct question is 80.7 percent, a difference of 23.4 percentage points. Similarly, the middle two bars in Figure 1 indicate that the estimated share of individuals who reported that they are comfortable working closely with a gay co-worker is 66.9 percent, whereas the share of individuals who reported that they are comfortable working closely with a gay co-worker in the direct question is 82.0 percent, a difference of 15.1 percentage points. Finally, the right-most two bars in Figure 1 indicate that the estimated share of individuals who reported that they are comfortable with a gay cashier is 65.7 percent, whereas the share of individuals who reported that they are comfortable with a gay cashier in the direct question is 87.4 percent, a difference of 21.7 percentage points. For context, the social desirability bias estimated in this context is somewhat higher than the average bias detected in previous list experiment meta-analyses (Blair, Coppock, and Moor 2020; Ehler, Wolter, and Junkermann 2021; Li and Van den Noortgate 2022) but is similar to that in other studies measuring attitudes toward sexism and diversity, equity, and inclusion (DEI) policies in the workplace (Boring and Delfgaauw 2024).

The next set of results involves investigating these differences via regression analyses. Specifically, the double list experiment design makes it possible to estimate the following regression model:

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \mu_i \qquad (1)$$

where T_i takes a value of 1 if the list includes the key statement (i.e., long list) and 0 otherwise. X_i is a vector of control variables that includes demographic controls (i.e., participant age, sex, sexual orientation, and current region-commune of residence), socioeconomic controls (i.e., participant educational level, employment status, income, religious views, political affiliation, and beliefs about the general level of comfort among the Chilean population with the three key statements), and a set of additional controls (i.e., whether at least one child less than 18 years of age lives in the participant's household, the total number of people living in the participant's

household, marital status, and day-of-week and week-of-sample indicators that represent the day of the week on which the participant started the experiment and the number of weeks since sample collection started).

Model (1) is estimated via ordinary least squares (OLS) separately for each list and for each key statement. The results are reported in Table 2. Columns 1-4 present the results from List A, whereas Columns 5-8 present the results from List B. Panel A shows the estimated share of individuals agreeing with the supervisor preference key statement, Panel B shows the estimated share of individuals agreeing with the co-worker preference key statement, and Panel C shows the estimated share of individuals agreeing with the customer preference key statement. In addition to these estimates, each panel reports the estimated bias calculated from the difference between each coefficient and the estimate obtained from the corresponding direct question.

As shown in Table 2, the estimates of $\hat{\beta}_1$, that is, the estimated fraction of the sample with the corresponding key attribute, are largely consistent and robust to the inclusion of a battery of controls. Additionally, all estimates are substantially smaller than the corresponding levels elicited from the direct questions. The estimated social desirability bias ranges from 15 to 35 percentage points for supervisor discomfort, from 7 to 27 percentage points for co-worker discomfort, and from 19 to 26 percentage points for customer discomfort.

3.2 Population beliefs

The next set of results focuses on the respondents' beliefs about attitudes among the general Chilean population toward supervising, working with, and buying from gay individuals. Figure 2 presents these results. Specifically, the respondents were asked to report their estimates in response to the following statements:

> "In the Chilean adult population, I think that approximately __ out of every 100 people would feel comfortable [supervising a gay employee] (Panel A) / [working closely with a gay co-worker] (Panel B) / [having a cashier at the supermarket who is gay] (Panel C)."

This elicitation of beliefs was designed to map directly onto the three supervisor, co-worker, and customer preference key statements introduced in the list experiments.

The histograms in Panels A and B roughly delineate normal distributions: most respondents believed that population-level comfort with gay employees and co-workers falls between 50 and 60 percent. On the other hand, Panel C displays a left-skewed distribution: the median is higher than the mean, and a substantial fraction of respondents believed that comfort with a gay cashier falls between 90 and 100 percent.[12]

In line with previous findings (Aksoy, Carpenter, and Sansone 2025), both the mean and median beliefs about comfort among the general population in all three panels are below the estimates from the list experiments of the share of the population comfortable with the corresponding key statements. In the figure, this comparison is presented in the box plot below

[12] Weighted kernel density versions are presented in Figure D1.

each histogram, where the white vertical line "|" indicates the median, the white "+" symbol indicates the mean, and the black "x" symbol indicates the share estimated from the list experiment. This result indicates that individuals underestimated the level of comfort with and support for gay workers in the labor market, especially in regard to supervising gay employees and having gay co-workers.

3.3 Heterogeneity analysis

This section focuses on estimating and presenting the results on the heterogeneous effects regarding the key findings. Based on insights from the previous literature (Badgett et al. 2024; Aksoy, Carpenter, and Sansone 2025), several demographic, socioeconomic, and other observable characteristics may moderate the effects presented thus far. Most of these observables are introduced as controls in the regression models presented in Table 2, e.g., participants' sex, age, sexual orientation, educational level, employment status, income, religious views, political affiliation, beliefs about the general population, marital status, locality, and children. Table 3 introduces the heterogeneous effects of each of these independent variables via an estimation method designed for double list experiments by Tsai (2019).

These heterogeneity analyses yield several insights. Young and middle-aged adults had higher shares of support for gay employees and co-workers than did older groups. Women had higher levels of comfort with gay co-workers and cashiers than did men. Compared to sexual minority individuals, heterosexual respondents had lower levels of comfort with gay individuals in all three roles, and the difference was statistically significant for the customer statement. Higher-income and non-college-educated individuals reported lower levels of comfort as co-workers and as customers of gay men, respectively.

On the other hand, employed participants were more comfortable with gay co-workers, while respondents with managerial experience reported higher levels of support across all three levels. It is worth highlighting these two sets of results, as employed individuals and those with managerial roles are those most likely to interact with gay individuals in the labor market and those for whom the two key statements were the most realistic.

Participants from outside the Metropolitan Region of Santiago reported lower comfort with gay co-workers, and those who lean politically left reported higher levels of comfort as supervisors. Those who did not identify with a religion were significantly more likely to feel comfortable with gay workers in any of the three roles than were those who declared a religious affiliation. Finally, holding more optimistic beliefs about the general population's attitudes toward gay people was related to higher reported levels of comfort.[13]

[13] Figures D2-D9 illustrate the results of univariate heterogeneity analyses, while Table D2 presents the results of multivariate analyses with additional controls. Notably, the univariate heterogeneity analyses estimate lower or no social desirability bias from respondents who identified as LGBTQ+ (Figure D3), those who did not report a religion (Figure D5), those who declared having a lesbian, gay, or bisexual relative or friend (Figure D8), and those who answered negatively to all direct questions about comfort with sexual minority individuals (Figure D9).

3.4 Comparison with other attitudes

This section focuses on comparing the answers to the direct key sensitive questions to the answers to the direct questions on attitudes toward sexual minority individuals in other contexts or attitudes regarding other sensitive topics. Specifically, in addition to the three supervisor, co-worker, and customer preference direct questions, the respondents were surveyed on whether they would feel comfortable:

- having a gay boss / dentist / real estate agent / taxi driver / waiter / neighbor.
- having to work closely with a lesbian co-worker.
- having to work closely with an indigenous co-worker.
- supervising several employees.

These questions were designed to test whether attitudes tend to be context specific and/or change according to the duration, frequency, and intimacy of the interactions, as well as for comparison with other minority groups. Figure 3 presents a summary of the findings. Overall, the majority of respondents reported that they would feel comfortable in all the scenarios presented. In fact, 66 percent of respondents answered "yes" to all the direct questions. This result is in line with Chile's recent history of progressiveness in social issues.

However, some variation emerged across statements. The highest share of comfort reported (92 percent of respondents) resulted from the indigenous co-worker statement. On the other hand, the lowest share of agreement (81 percent of respondents) was prompted by the statement on supervising a gay employee. In fact, the idea of supervising a gay employee elicited the lowest share of individuals feeling comfortable consistently throughout the study, regardless of the elicitation method. The statement about gay co-workers elicited more support but slightly less than the statement about lesbian co-workers and less than the statements about interactions with gay people in shorter or more distant contexts. Notably, the discomfort with supervising gay employees did not stem from a general uneasiness with holding a managerial role: most participants reported feeling comfortable supervising multiple employees, considerably more so than when the subordinate was gay.

The averages reported in Figure 3 also highlight that shorter interactions with gay individuals, such as checking out at the supermarket with a gay cashier, having a gay realtor, having a gay taxi driver or having a gay waiter, resulted in higher levels of individuals reporting that they felt comfortable. In line with previous studies highlighting the role of proximity in driving discrimination (Smerdon, Pearson, and Albrecht 2025), distant interactions – such as having a gay neighbor, as opposed to closely working with a gay person – also elicited a high share of individuals expressing comfort. These findings suggest that the level of comfort tends to decrease as the perceived intimacy or frequency of interaction increases. While interactions in casual service roles often result in high acceptance, situations that involve sustained or direct collaboration, such as working closely or supervising, appear to evoke more discomfort. This finding highlights potentially complex dynamics in attitudes, where proximity and role expectations seem to influence the reported comfort levels.

Having said that, it is important to emphasize that Figure 3 reports stated support: the findings may be different when correcting for social desirability bias. Indeed, the level of comfort with a gay cashier and the level of comfort with a gay co-worker, elicited from the list experiment, are closer in Figure 1. Overall, these findings suggest that the reported levels of comfort with gay individuals are affected by proximity/duration, role expectations, and – in the case of survey-based measures – social desirability bias.

3.5 Donation behavior

The results presented thus far have focused on the double list experiments and survey measures collected from an extensive, nationally representative sample that is considerably larger than those used in most list experiments in the literature. This section investigates the consistency of the main results after the sample is split based on the participants' choices in a real-stakes, incentive-compatible donation question.

Toward the end of the survey, all participants were entered into a raffle for an additional payment of approximately 100 USD. They were informed that the probability of winning this payment was uniform for all participants who completed the study and that it was not affected by their behavior. Following the introduction of the lottery but before the announcement of the results, individuals were asked whether they would donate any amount of their earnings to one of two local LGBTQ-related nonprofit organizations if they were to be randomly selected to win this prize.[14] The respondents were randomly assigned one of the two nonprofit organizations and were provided with their mission statements to standardize the minimum amount of information that all participants had on the organizations. The participants were informed that the wishes of the winners would be honored. That is, any amount that they chose not to donate was theirs to keep (added to their compensation for participation), and any amount that they chose to donate would be sent to the corresponding organization by the partner survey company, Datavoz, after completion of the experiment.

Data across the two nonprofit organizations were pooled, and 1,683 (42.1 percent of the sample) chose not to donate any of their earnings if they were to win the raffle. On the other hand, 2,317 (57.9 percent of the sample) voluntarily chose to donate some or all of their earnings (Figure D10).[15] This donation behavior correlates with the preferences elicited throughout the survey. Figure 4 presents the main list experiment results when the sample is split into donors and non-donors. Among donors in all three capacities (i.e., supervisor, co-workers, and customer), the level of support was higher, and the estimated social desirability bias was lower than that among respondents who chose to retain the entirety of their monetary prizes. In particular, social desirability bias was substantially and (marginally) significantly

[14] Appendix B.1 provides more information about the local nonprofit organizations.

[15] Donation behavior is qualitatively similar across nonprofit organizations. Specifically, if they were to win the raffle, the weighted share of individuals who chose to not donate a positive amount was 43.8 percent if the respondents saw Movilh as their nonprofit organization and 42.7 percent if the respondents saw Iguales as their nonprofit organizations (Figure D10).

different between donors and non-donors for the co-worker statement. These results are largely consistent across both nonprofit organizations (Figure D11).[16]

4. Conclusion

This study provides evidence of distaste toward gay individuals in the workplace. Many respondents in Chile tend to express comfort when asked directly about their views on supervising, working closely, and buying from gay individuals. However, the estimates from the double list experiments indicate that many of these respondents are misreporting their attitudes.

These results indicate that distaste can drive taste-based discrimination against sexual minority individuals in the workplace. Although the level of discomfort and the estimated social desirability bias are the highest when the respondents consider supervising a gay employee, the other list experiment findings suggest that distaste is prevalent even among co-workers and customers. From a methodological perspective, these findings raise concerns that traditional polling and survey methods may fail to accurately capture public attitudes toward sexual minority individuals.

In line with group identity and contact theory, the level of support is the highest among LGBTQ+ respondents and those who know someone who is gay. Women, young people, employed individuals, and educated individuals are also more supportive. Similarly, higher comfort levels are found among respondents with managerial experience. Importantly, despite these differences and the large estimated social desirability bias, the majority of respondents in all categories express comfort with interacting with gay individuals in the workplace, thus indicating broad societal acceptance levels.

When these attitudes are compared to those toward sexual minority individuals in other contexts or toward other minority individuals, the evidence suggests that individuals report a higher level of acceptance for indigenous people in the workplace than for sexual minority individuals. The reported levels of support for gay and lesbian individuals in the workplace are comparable, while support for a gay cashier aligns closely with that in other scenarios. However, in regard to supervising gay employees or working alongside a gay co-worker, the comfort levels are lower.

Additional survey results reveal that the respondents underestimate the level of support for gay employees and co-workers within the Chilean population, suggesting the potential for an informational campaign to address these misconceptions (Bursztyn, González, and Yanagizawa-Drott 2020). Finally, attitudinal differences correlate with real-stakes donation behavior: among individuals who chose not to donate any amount to the LGBTQ-related nonprofit organizations, the degree of misreporting is higher, and the level of comfort is lower.

This experiment focused on attitudes toward sexual minority individuals. With few exceptions (Aksoy, Carpenter, and Sansone 2025), very little is known about attitudes toward gender minority individuals such as transgender and non-binary individuals. Future research could explore how the reported attitudes toward gay individuals in the workplace compare to attitudes

[16] The results of additional heterogeneity analyses are presented in Figures D12-D13, which document a gradient in attitudes by donation size: those in the top half or top quartile of donors display more supportive attitudes and lower misreporting than do those in the bottom segments.

in other contexts such as sports, media and entertainment, or educational settings. Furthermore, as noted by (Blair, Coppock, and Moor 2020), list experiments address bias in measures of explicit attitudes: more research is needed to measure implicit attitudes toward sexual and gender minority individuals, as recently done for racial minority individuals in a similar context (Duryea et al. 2025). Finally, future studies could advance this research agenda and complement these results by using measures of discomfort with sexual and gender minority individuals that are quantifiable in dollar terms, for example, by measuring actions or behaviors that reveal a willingness to pay to avoid a gay employee, co-worker, or salesperson, among others.

From a policy perspective, these findings underscore that stated support for minority individuals – often influenced by a desire to conform to social norms or legal expectations – may not reflect individuals' true attitudes or behaviors. This mismatch helps explain why large socioeconomic and health disparities by sexual orientation persist, despite growing public expressions of acceptance for LGBTQ+ individuals in opinion surveys. Real and sustained reductions in these disparities may require deeper cultural change and a meaningful decrease in both conscious and unconscious biases.

Moreover, identifying the extent of misreporting around comfort with minority individuals – even in relatively progressive countries such as Chile – can reinforce the ongoing need for anti-discrimination policies, particularly when high levels of stated tolerance mask underlying prejudice. Similarly, if a small group of individuals with strongly negative views can influence aggregate wage gaps and disparities – as suggested by (Sterkens et al. 2022) and formalized in search models (Black 1995; Maloney and Neumark 2025) – then socioeconomic gaps by sexual orientation may persist even in contexts where the majority of the population holds positive attitudes toward sexual minority individuals in the workplace.

References

Abbate, Nicolás, Inés Berniell, Joaquín Coleff, Luis Laguinge, Margarita Machelett, Mariana Marchionni, Julián Pedrazzi, and María Florencia Pinto. 2024. "Discrimination against Gay and Transgender People in Latin America: A Correspondence Study in the Rental Housing Market." *Labour Economics* April: 102486.

Agüero, Jorge M., and Veronica Frisancho. 2022. "Measuring Violence against Women with Experimental Methods." *Economic Development and Cultural Change* 70 (4): 1565–90.

Aksoy, Billur, Christopher S. Carpenter, and Dario Sansone. 2025. "Understanding Labor Market Discrimination Against Transgender People: Evidence from a Double List Experiment and a Survey." *Management Science* 71 (1): 659–77.

Aksoy, Billur, Ian Chadd, and Boon Han Koh. 2023. "Sexual Identity, Gender, and Anticipated Discrimination in Prosocial Behavior." *European Economic Review* 154: 104427.

Arrow, Kenneth. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets*. Princeton: Princeton University Press.

Badgett, M.V. Lee, Christopher S. Carpenter, Maxine J. Lee, and Dario Sansone. 2024. "A Review of the Economics of Sexual Orientation and Gender Identity." *Journal of Economic Literature* 62 (3): 948–94.

———. 2025. "Effects of Legal Access to Same-Sex Marriage." *Journal of Policy Analysis and Management* 41 (1): 266294.

Badgett, M.V. Lee, Christopher S. Carpenter, and Dario Sansone. 2021. "LGBTQ Economics." *Journal of Economic Perspectives* 35 (2): 141–70.

Bar, Revital, and Asaf Zussman. 2017. "Customer Discrimination: Evidence from Israel." *Journal of Labor Economics* 35 (4): 1031–59.

Becker, Gary S. 1971. *The Economics of Discrimination*. 2nd ed. Chicago, IL: University of Chicago Press.

Bertrand, Marianne, and Esther Duflo. 2017. "Field Experiments on Discrimination." *Handbook of Economic Field Experiments* 1 (January): 309–93.

Black, Dan A. 1995. "Discrimination in an Equilibrium Search Model." *Journal of Labor Economics* 13 (2): 309–34.

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." *American Political Science Review* 114 (4): 1297–1315.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789–865.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2025. "Inaccurate Statistical Discrimination: An Identification Problem." *The Review of Economics and Statistics* 107 (3): 605–20.

Bohren, J. Aislinn, Peter Hull, and Alex Imas. 2025. "Systemic Discrimination: Theory and Measurement." *Quarterly Journal of Economics* Accepted.

Boring, Anne, and Josse Delfgaauw. 2024. "Social Desirability Bias in Attitudes towards Sexism and DEI Policies in the Workplace." *Journal of Economic Behavior & Organization* 225: 465–82.

Borm, Hannah Van, Marlot Dhoop, Allien Van Acker, and Stijn Baert. 2020. "What Does Someone's Gender Identity Signal to Employers?" *International Journal of Manpower* 41 (6): 753–77.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110 (10): 2997–3029.

Button, Patrick. 2018. "Expanding Employment Discrimination Protections for Individuals with Disabilities: Evidence from California." *ILR Review* 71 (2): 365–93.

Castanho Silva, Bruno, Mario Fuks, and Eduardo Ryô Tamaki. 2022. "So Thin It's Almost Invisible: Populist Attitudes and Voting Behavior in Brazil." *Electoral Studies* 75 (February): 102434.

Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban. 2021. "Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods." *Journal of Development Economics* 148 (January): 102582.

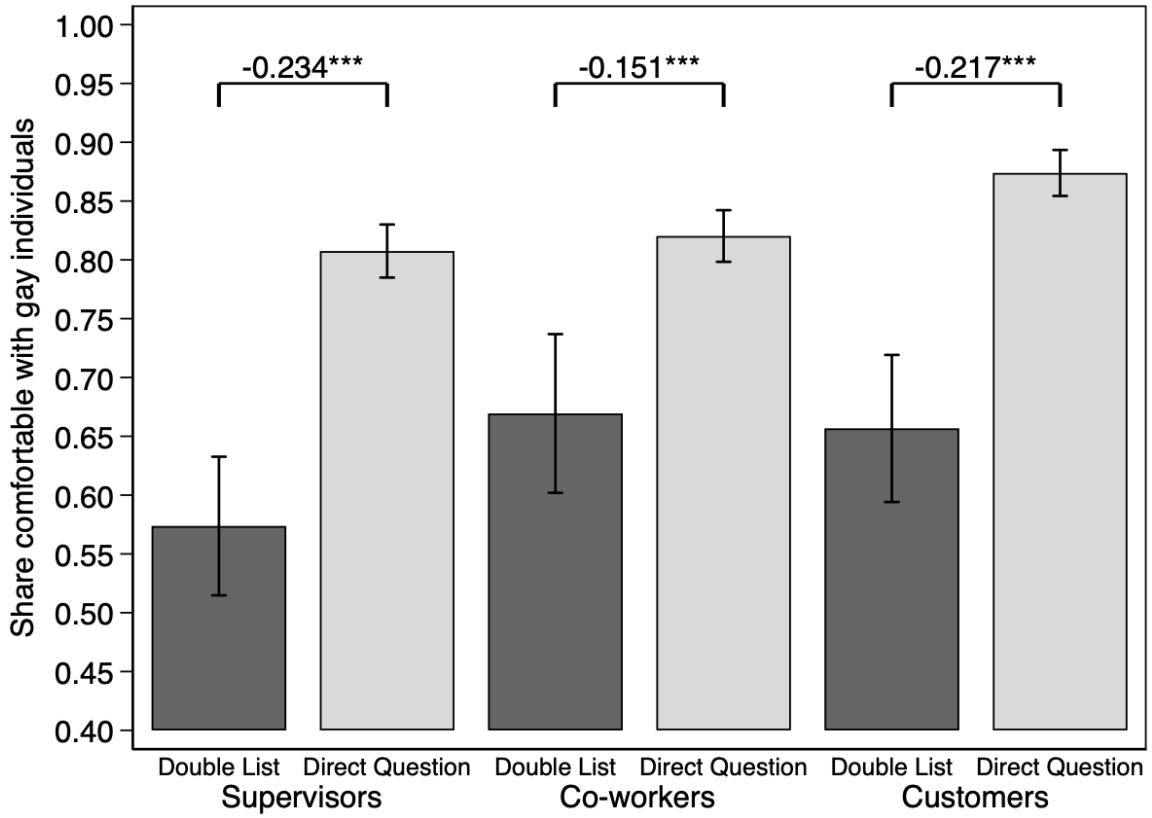
Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson. 2017. "The Size of the LGBT

Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated." *Management Science* 63 (10): 3168–86.
Coutts, Elisabeth, and Ben Jann. 2011. "Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT)." *Sociological Methods & Research* 40 (1): 169–93.
Dolado, Juan J., Luigi Minale, and Airam Guerra. 2023. "Uncovering the Roots of Obesity-Based Wage Discrimination: The Role of Job Characteristics." *Labour Economics* 85 (July): 102425.
Donohue, John J. III, and James J. Heckman. 1991. "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Black." *Journal of Economic Literature* 29 (4): 1603–43.
Droitcour, Judith, Rachel A. Caspar, Michael L. Hubbard, Teresa L. Parsley, Wendy Visscher, and Trena M. Ezzati. 1991. "The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application." In *Measurement Errors in Surveys*, edited by P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman, 185–210. New York, NY: John Wiley & Sons, Ltd.
Duryea, Suzanne, Jaime Millán-Quijano, Judith Morrison, and Yanira Oviedo. 2025. "Measuring Racial Bias in Employment Services in Colombia." *Journal of Development Economics* 174: 103435.
Ehler, Ingmar, Felix Wolter, and Justus Junkermann. 2021. "Sensitive Questions in Surveys: A Comprehensive Meta-Analysis of Experimental Survey Studies on the Performance of the Item Count Technique." *Public Opinion Quarterly* 85 (1): 6–27.
Friese, Malte, Colin Tucker Smith, Thomas Plischke, Matthias Bluemke, and Brian A. Nosek. 2012. "Do Implicit Attitudes Predict Actual Voting Behavior Particularly for Undecided Voters?" *PLOS ONE* 7 (8): e44130.
Gervais, Will M., and Maxine B. Najle. 2018. "How Many Atheists Are There?" *Social Psychological and Personality Science* 9 (1): 3–10.
Glasman, Laura R., and Dolores Albarracín. 2006. "Forming Attitudes That Predict Future Behavior: A Meta-Analysis of the Attitude-Behavior Relation." *Psychological Bulletin* 132 (5): 778–822.
Glynn, Adam N. 2013. "What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment." *Public Opinion Quarterly* 77 (S1): 159–72.
Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119.
———. 2021. *Career and Family*. Princeton, NJ: Princeton University Press.
Gutiérrez, Emilio, and Adrian Rubli. 2024a. "LGBT+ Persons and Homophobia Prevalence across Job Sectors: Survey Evidence from Mexico." *Labour Economics* 87 (April): 102500.
———. 2024b. "LGBTQ+ Individuals in the Mexican Labor Market: Queerphobia, Sorting, and Observable Outcomes." *Working Paper*.
Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1): 3–40.
Ham, Andres, Angela Guarin, and Juanita Ruiz. 2024. "How Accurately Are Household Surveys Measuring the LGBT Population in Colombia? Evidence from a List Experiment." *Labour Economics* 87 (April): 102503.
Holbrook, Allyson L., and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74 (1): 37–67.
Jackson, Chris. 2023. "LGBT+ Pride 2023: A 30-Country Ipsos Global Advisor Survey." *Ipsos*.
Jamison, Julian, Dean S. Karlan, and Pia Raffler. 2013. "Mixed Method Evaluation of a Passive Health Sexual Information Texting Service in Uganda." *Information Technologies & International Development* 9 (3): 1–28.
Jonge, Chad P. Kiewiet de. 2015. "Who Lies About Electoral Gifts? Experimental Evidence from Latin America." *Public Opinion Quarterly* 79 (3): 710–39.
Kelley, Erin Munro, Gregory Lane, Matthew Pecenco, and Edward Rubin. 2025. "Customer Discrimination in the Workplace: Evidence from Online Sales." *Journal of Labor Economics* Accepted (September).
Kessler, Judd B., Corinne Low, and Xiaoyue Shan. 2025. "Lowering the Playing Field: Discrimination through Sequential Spillover Effects." *The Review of Economics and Statistics* Accepted (February).

Klawitter, Marieka. 2011. “Multilevel Analysis of the Effects of Antidiscrimination Policies on Earnings by Sexual Orientation.” *Journal of Policy Analysis and Management* 30 (2): 334–58.
Klawitter, Marieka, and Victor Flatt. 1998. “The Effects of State and Local Antidiscrimination Policies on Earnings for Gays and Lesbians.” *Journal of Policy Analysis and Management* 17 (4): 658–86.
Lax, Jeffrey R., Justin H. Phillips, and Alissa F. Stollwerk. 2016. “Are Survey Respondents Lying about Their Support for Same-Sex Marriage? Lessons from a List Experiment.” *Public Opinion Quarterly* 80 (2): 510–33.
Li, Jiayuan, and Wim Van den Noortgate. 2022. “A Meta-Analysis of the Relative Effectiveness of the Item Count Technique Compared to Direct Questioning.” *Sociological Methods & Research* 51 (2): 760–99.
Lippens, Louis, Siel Vermeiren, and Stijn Baert. 2023. “The State of Hiring Discrimination: A Meta-Analysis of (Almost) All Recent Correspondence Experiments.” *European Economic Review* January: 104315.
Maloney, Molly, and David Neumark. 2025. “Does the Gender Wage Gap Actually Reflect Taste Discrimination Against Women?” *NBER Working Paper* 33405.
McKenzie, David, and Melissa Siegel. 2013. “Eliciting Illegal Migration Rates through List Randomization.” *Migration Studies* 1 (3): 276–91.
Meyer, Ilan H. 1995. “Minority Stress and Mental Health in Gay Men.” *Journal of Health and Social Behavior* 36 (1): 38.
Miller, Judith Droitcour. 1984. “A New Survey Technique for Studying Deviant Behavior.” *Dissertation*, 1–198.
Montoya, Ana María, Eric Parrado, Ales Solís, and Raimundo Undurraga. 2025. “Bad Taste: Gender Discrimination in Consumer Lending.” *Journal of Political Economy: Microeconomics* Accepted (November).
Muñoz, Ercio A., Melanie Saavedra, and Dario Sansone. 2024. “Socio-Economic Disparities by Sexual Orientation and Gender Identity in Mexico.” *IDB Working Paper*, 1–21.
———. 2025. “The Lives of Intersex People: Socioeconomic and Health Disparities in Mexico.” *PNAS Nexus* 4 (5): pgaf126.
Muñoz, Ercio A., and Dario Sansone. 2024. “Matching Patterns among Same-Sex and Different-Sex Couples in Latin America.” *AEA Papers and Proceedings* 114.
Muñoz, Ercio A., Dario Sansone, and Mayte Ysique. 2024. “Socio-Economic Disparities in Latin America among Same-Sex and Different-Sex Couples.” *IDB Working Paper*, 1–71.
NASEM. 2020. “Understanding the Well-Being of LGBTQI+ Populations.” *National Academies of Sciences, Engineering, and Medicine*, 1–436.
Nettuno, Laura. 2024. “Gender Identity, Labor Market Outcomes, and Socioeconomic Status: Evidence from Chile.” *Labour Economics* 87: 102487.
Neumark, David. 2018. “Experimental Research on Labor Market Discrimination.” *Journal of Economic Literature* 56 (3): 799–866.
Neumark, David, Ian Burn, Patrick Button, and Nanneh Chehras. 2019. “Do State Laws Protecting Older Workers from Discrimination Reduce Age Discrimination in Hiring? Evidence from a Field Experiment.” *Journal of Law and Economics* 62 (2): 373–402.
Neumark, David, and Wendy A. Stock. 2006. “The Labor Market Effects of Sex and Race Discrimination Laws.” *Economic Inquiry* 44 (3): 385–419.
Onuchic, Paula. 2024. “Recent Contributions to Theories of Discrimination.” *Working Paper*.
Osman, Adam, Jamin D. Speer, and Andrew Weaver. 2025. “Discrimination Against Women in Hiring.” *Economic Development* 73 (2): 781–809.
Palacios, Cristóbal. 2024. “Ley de Presupuestos: Cámara Prohíbe Al Minsal Financiar Terapias Hormonales Para Tratar Disforia de Género En Menores.” *La Tercera* November (14).
Plug, Erik, Dinand Webbink, and Nick Martin. 2014. “Sexual Orientation, Prejudice, and Segregation.” *Journal of Labor Economics* 32 (1): 123–59.
Raghavarao, D., and W. T. Federer. 1979. “Block Total Response as an Alternative to the Randomized Response Method in Surveys.” *Journal of the Royal Statistical Society: Series B (Methodological)* 41 (1): 40–45.
Rayburn, Nadine Recker, Mitchell Earleywine, and Gerald C. Davison. 2003. “An Investigation of Base Rates of Anti-Gay Hate Crimes Using the Unmatched-Count Technique.” *Journal of Aggression,*

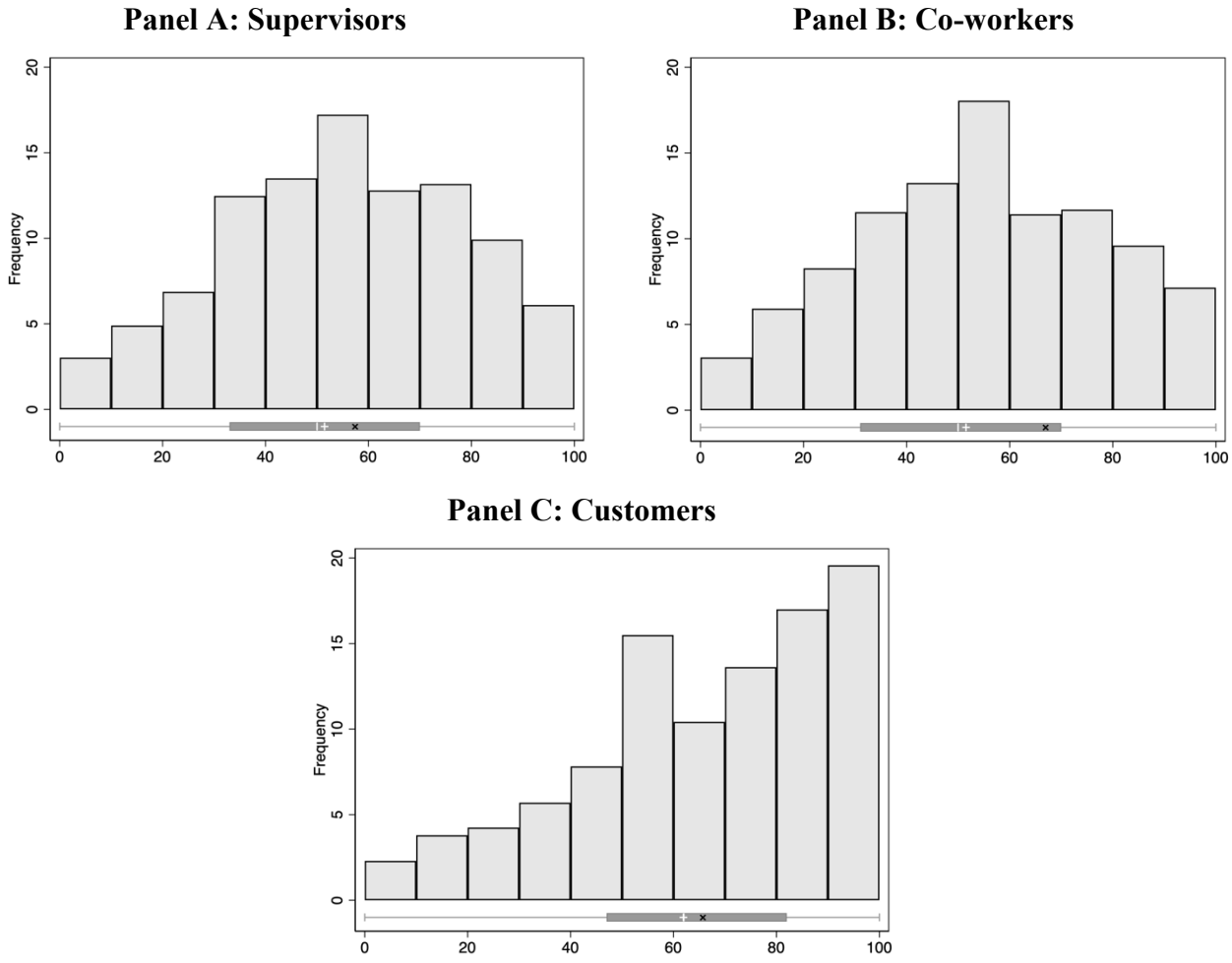
Maltreatment & Trauma 6 (2): 137–52.
Robertson, Ronald E., Felix W. Tran, Lauren N. Lewark, and Robert Epstein. 2018. "Estimates of Non-Heterosexual Prevalence: The Roles of Anonymity and Privacy in Survey Methodology." *Archives of Sexual Behavior* 47 (4): 1069–84.
Sansone, Dario. 2019. "Pink Work: Same-Sex Marriage, Employment and Discrimination." *Journal of Public Economics* 180 (December): 104086.
Small, Mario L., and Devah Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives* 34 (2): 49–67.
Smerdon, David, Samuel Pearson, and Sabina Albrecht. 2025. "Physical Proximity Drives Gay Discrimination in the Gig Economy." *Proceedings of the National Academy of Sciences* 122 (11): e2412362122.
Stephens-Davidowitz, Seth. 2014. "The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data." *Journal of Public Economics* 118: 26–40.
Sterkens, Philippe, Axana Dalle, Joey Wuyts, Ines Pauwels, Hellen Durinck, and Stijn Baert. 2022. "Homosexuality's Signalling Function in Job Candidate Screening: Why Gay Is (Mostly) Ok." *Working Paper*.
Streb, Matthew J., Barbara Burrell, Brian Frederick, and Michael A. Genovese. 2008. "Social Desirability Effects and Support for a Female American President." *Public Opinion Quarterly* 72 (1): 76–89.
Tampellini, João. 2024. "Latin American Pride: Labor Market Outcomes of Sexual Minorities in Brazil." *Journal of Development Economics*, no. 167: 103239.
Webb, Duncan. 2025. "Silence to Solidarity: How Communication About a Minority Affects Discrimination." *Working Paper*.
World Bank. 2022. "Individuals Using the Internet (% of Population) - Chile." *World Bank DataBank*. https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=CL&most_recent_value_desc=true.

Figure 1. List Experiments on Supervisors', Co-workers', and Customers' Attitudes toward Gay Individuals



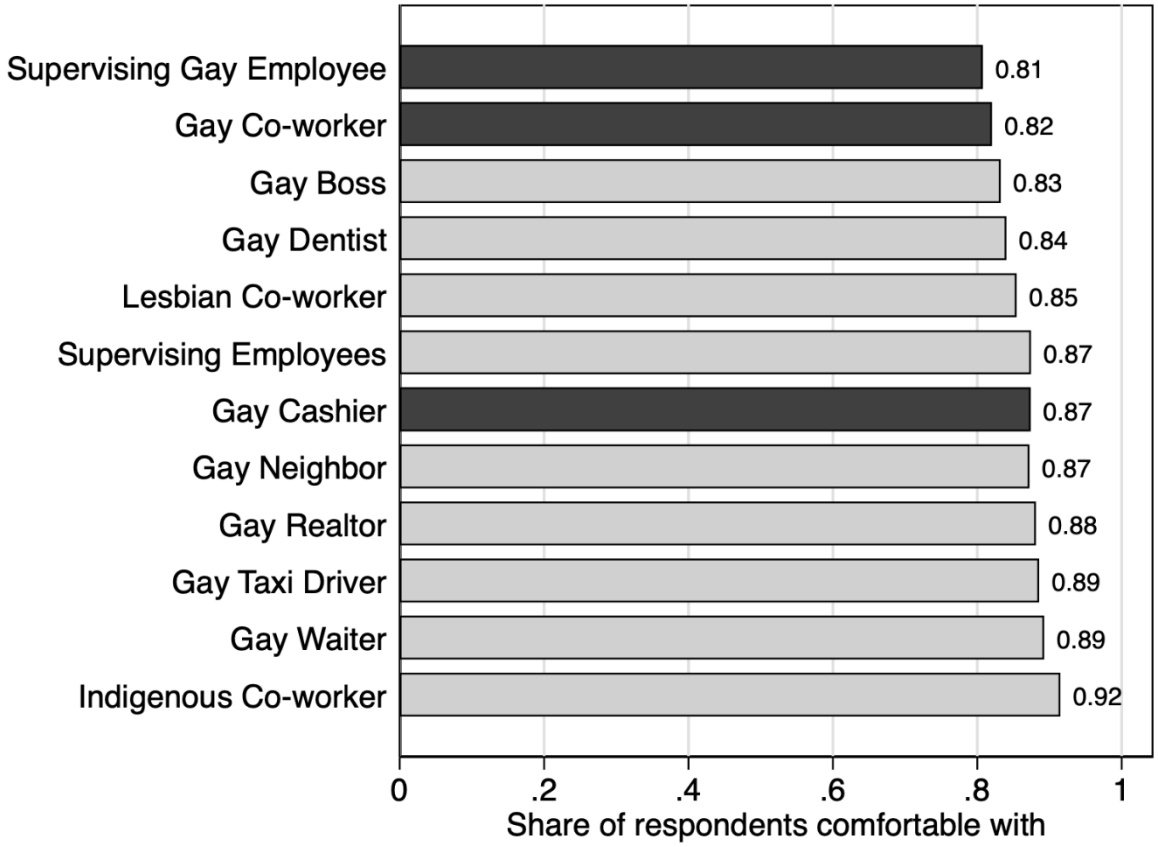
Notes: Weighted statistics. 95-percent confidence intervals are reported via the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. Supervisor key statement: "I would feel comfortable supervising a gay employee." Co-worker key statement: "I would feel comfortable working closely with a gay co-worker." Customer key statement: "I would feel comfortable having a cashier at the supermarket who is gay." Number of observations: 4,000. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Figure 2. Perceptions of General Views on Attitudes toward Gay People



Notes: Weighted statistics (and unweighted histogram). The original survey questions are as follows: "In the adult Chilean population, I think that approximately __ out of every 100 people would feel comfortable [supervising a gay employee] (Panel A) / [working closely with a gay co-worker] (Panel B) / [having a cashier at the supermarket who is gay] (Panel C)." The box plot below each histogram reports the minimum and maximum values and the 25th and 75th percentiles, as well as the mean and median. Within each box plot, the white vertical line "|" indicates the median, and the white "+" symbol indicates the mean. The black "x" symbol in Panel A indicates the share of the sample that is comfortable supervising a gay employee, estimated from the double list experiment in Figure 1; in Panel B, it indicates the share of the sample that is comfortable with a gay co-worker, estimated from the corresponding double list experiment; and in Panel C, it indicates the estimated share of the sample that is comfortable with a gay cashier. Weighted kernel densities are presented in Figure D1. Number of observations: 4,000.

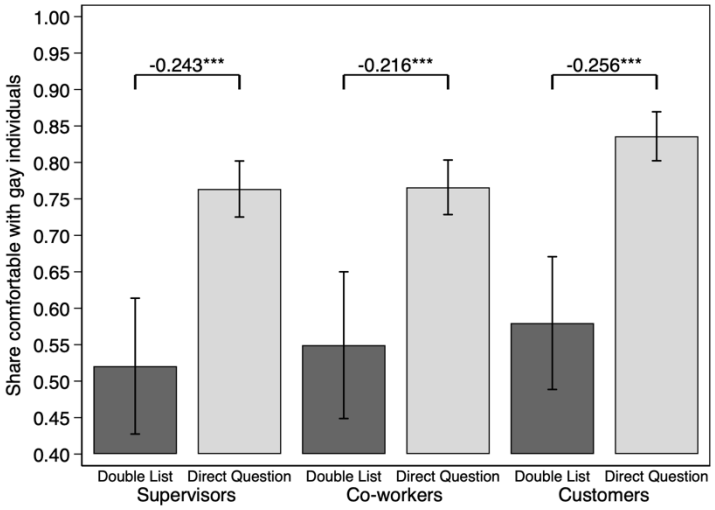
Figure 3. Comparison of Views on Gay Individuals in Multiple Contexts and on Other Minority Groups



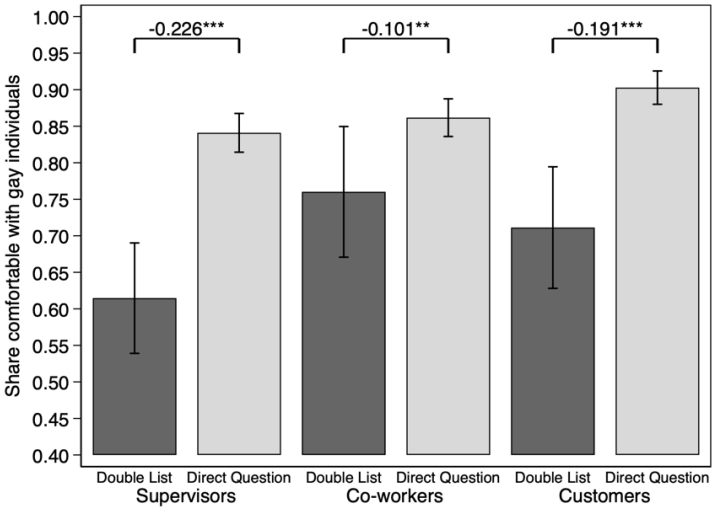
Notes: Weighted statistics. The bars in darker gray represent the responses to the three direct questions corresponding to the employee, co-worker, and customer preference key statements introduced in the list experiments (see Figure 1). The lighter gray bars correspond to the following direct questions (listed in the order in which they appear in the figure): "Would you feel comfortable… [having a gay boss] / [having a gay dentist] / [having to work closely with a lesbian co-worker] / [supervising several employees] / [having a gay neighbor] / [having a gay real estate agent] / [having a gay taxi driver] / [being served by a gay waiter] / [having to work closely with an indigenous co-worker]?". Number of observations: 4,000.

Figure 4. List Experiments on Supervisors', Co-workers', and Customers' Attitudes toward Gay Individuals by Donor Status

Panel A: Nondonors



Panel B: Donors



Notes: Weighted statistics. 95-percent confidence intervals are reported via the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Panel A: The sample is restricted to individuals who chose to not donate any amount of a raffled monetary prize to one of two local LGBTQ-related nonprofit organizations (number of observations: 1,683). Panel B: The sample is restricted to individuals who chose to donate a positive amount of a raffled monetary prize to one of two local LGBTQ-related nonprofit organizations (number of observations: 2,317). *p < 0.10; **p < 0.05; ***p < 0.01.

Table 1: List Experiment Example

Short list	**Long list**
I own a car or motorcycle. I have a lot of confidence in political parties. I think the military should work with the police to fight crime. I believe that half of the legislators in Congress should be women.	I own a car or motorcycle. I have a lot of confidence in political parties. I think the military should work with the police to fight crime. I believe that half of the legislators in Congress should be women. I would feel comfortable having a cashier at the supermarket who is gay. *[key statement]*

Notes: The order of the statements within each list was randomized at the subject level. For the full set of lists, as well as the original Spanish questionnaire, see Appendix E.

Table 2: List Experiments on Supervisors', Co-workers', and Customers' Attitudes toward Gay Individuals

	List A				List B			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Supervisors								
Subject saw list with key statement	0.676	0.664	0.640	0.667	0.472	0.480	0.496	0.504
	(0.054)	(0.044)	(0.043)	(0.044)	(0.050)	(0.047)	(0.050)	(0.049)
R^2	0.152	0.234	0.220	0.242	0.076	0.133	0.156	0.207
Estimated bias	15.075	16.275	18.675	15.975	35.475	34.675	33.075	32.275
Panel B: Co-workers								
Subject saw list with key statement	0.772	0.754	0.740	0.773	0.567	0.580	0.594	0.611
	(0.054)	(0.048)	(0.048)	(0.050)	(0.046)	(0.045)	(0.047)	(0.046)
R^2	0.174	0.228	0.247	0.283	0.115	0.162	0.191	0.238
Estimated bias	6.575	8.375	9.775	6.475	27.075	25.775	24.375	22.675
Panel C: Customers								
Subject saw list with key statement	0.685	0.684	0.671	0.709	0.628	0.637	0.645	0.656
	(0.054)	(0.047)	(0.047)	(0.046)	(0.055)	(0.050)	(0.053)	(0.051)
R^2	0.140	0.202	0.205	0.268	0.107	0.192	0.217	0.267
Estimated bias	20.75	20.85	22.15	18.35	26.45	25.55	24.75	23.65
Controls for:								
Region-commune fixed effects		✓	✓	✓		✓	✓	✓
Demographic controls		✓	✓	✓		✓	✓	✓
Socioeconomic factors			✓	✓			✓	✓
Additional controls				✓				✓
Observations	4,000	3,947	3,514	3,293	4,000	3,947	3,514	3,293

Notes: Multivariate weighted analysis. OLS estimates with robust standard errors in parentheses. All estimates are highly significant at the 1-percent level. "Estimated bias" reports the differences (in percentage points) between the estimated percentage of participants who agreed with the key statement in each corresponding column and the estimate obtained from the corresponding direct question (as reported in Figure 1). Demographic controls include the participant age, sex, sexual orientation, and gender identity as well as indicators for ethnicity (i.e., African descent) and indigenous status. Region-commune fixed effects include the current region of residence or commune for those living in the Metropolitan Region of Santiago. Socioeconomic factors include the participant educational level, income, employment status, managerial experience, religious views, and political affiliation. Additional controls include whether at least one child less than 18 years of age lives in the participant's household, the number of people living in the participant's household, marital status, lesbian/gay comfort, an indicator for whether the participant knows a lesbian, gay, or bisexual person, and day-of-week and week-of-sample indicators. All variables are described in Appendix B.

Table 3: List Experiment on Supervisors', Co-workers', and Customers' Attitudes toward Gay Individuals – Heterogeneity Analyses

Interaction of treatment variable with:	Supervisors (1)	Co-workers (2)	Customers (3)
Age: 18–44	0.126*	0.125*	-0.049
	(0.068)	(0.072)	(0.066)
Race: African Descent	-0.345	-0.001	0.082
	(0.224)	(0.178)	(0.217)
Indigenous	0.027	0.159	-0.031
	(0.121)	(0.124)	(0.127)
Sex: Female	0.009	0.164**	0.203***
	(0.065)	(0.070)	(0.067)
Sexual orientation: Heterosexual	-0.178	-0.254	-0.388***
	(0.124)	(0.181)	(0.113)
Household income: More than $1,500	0.074	-0.171**	0.002
	(0.064)	(0.068)	(0.068)
Education: More than high school	0.042	0.066	0.144*
	(0.085)	(0.089)	(0.081)
Employment status: Employed	0.098	0.135*	-0.043
	(0.083)	(0.079)	(0.086)
Management Experience	0.147**	0.143*	0.171**
	(0.073)	(0.080)	(0.068)
Region: Outside metro	0.006	-0.119*	-0.062
	(0.059)	(0.065)	(0.061)
Political affiliation: Lean left	0.094	0.036	0.017
	(0.067)	(0.076)	(0.068)
Current religious affiliation: Not religious	0.125*	0.115	0.220***
	(0.070)	(0.078)	(0.073)
Belief: 50% or more comfortable supervising gay employees	0.109*		
	(0.064)		
Belief: 50% or more comfortable with gay co-workers		0.179***	
		(0.068)	
Belief: 50% or more comfortable with gay cashiers			0.050
			(0.072)
Constant	0.300*	0.498**	0.667***
	(0.164)	(0.223)	(0.170)
Observations	3,518	3,518	3,518

Notes: Heterogeneity weighted multivariate analysis. Robust standard errors are in parentheses. Coefficients were obtained via the Stata command kict ls (Tsai 2019) to perform least squares estimation for a double list experiment. The dependent variables are the reported number of true statements for the gay employee supervisor lists (Column 1), the gay co-worker lists (Column 2), and the gay cashier lists (Column 3). The treatment variable is an indicator equal to 1 for the first long list (List A) containing the corresponding key statement and the second short list (List B) and equal to 0 for the first short list (List A) and the second long list (List B). The sample size is lower than 4,000 since missing observations reflect cases in which the participants did not provide an answer or selected options such as "I prefer not to respond". *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Online Appendix for "Measuring the Sources of Taste-Based Discrimination Using List Experiments" (NOT MEANT FOR PUBLICATION)

Appendix A. Technical details

A.1 Ethics and pre-registration

At the beginning of the experiment, the respondents signed a consent form (reported in Appendix E.2). Only individuals older than 18 were allowed to take part in this study. This research was approved as exempt by the IRB at the University of Maryland (#2186752) and was approved by the Research Ethics Committee at the University of Exeter (#6475933).

The experiment and a pre-analysis plan were pre-registered on the American Economic Association's registry for randomized controlled trials on July 9, 2024, and published online on July 16, 2024 (AEARCTR-0013989): https://doi.org/10.1257/rct.13989-1.0.

A.2 Formal details on the list experiment technique

A.2.1 Mean comparisons in list experiments

The main list experiment analysis follows the standard estimation technique implemented in previous studies (Tsai 2019; Aksoy, Carpenter, and Sansone 2025). Suppose that there is a sample of *n* participants. Let T_i be the indicator variable equal to one if participant *i* sees the long list with the key sensitive item instead of the short list and 0 otherwise. Let S_i be participant *i*'s potential answer to the key statement, and let $R_{i,j}$ be participant *i*'s potential answer to the jth non-key statement (where j=4 in this application). Via the list in Table 1, $S_i = 1$ if participant *i* would be comfortable having a cashier at the supermarket who is gay and 0 otherwise. Similarly, for example, $R_{i,2} = 1$ if participant *i* thinks that the military should work with the police to fight crime and 0 otherwise. Notably, the researchers do not observe S_i or $R_{i,j}$. Instead, they observe the total number of statements that are true for participant *i*: $Y_i = T_i S_i + R_i$, where $R_i = \sum_{j=1}^{4} R_{i,j}$. Under certain assumptions discussed in Appendix A.5, the following difference-in-means estimator gives the estimated share of the population with the key attribute (i.e., $E(S_i)$).

$$E(S_i) = \frac{\sum_{i=1}^{n} Y_i T_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n} Y_i (1 - T_i)}{\sum_{i=1}^{n} (1 - T_i)} \quad (1)$$

A.2.2 Mean comparisons in double list experiments

To formally introduce the double list experiment technique, let Y_i^A and Y_i^B be the total number of items that are true for participant *i* in the two list experiments with the same key statement but different non-key statements (Lists A and B, respectively). The estimated share of the population with the key attribute is given by $E^{DL}(S_i)$.

$$E^{DL}(S_i) = \left[\left\{ \frac{\sum_{i=1}^{n} Y_i^A T_i}{\sum_{i=1}^{n} T_i} - \frac{\sum_{i=1}^{n} Y_i^A (1 - T_i)}{\sum_{i=1}^{n} (1 - T_i)} \right\} + \left\{ \frac{\sum_{i=1}^{n} Y_i^B (1 - T_i)}{\sum_{i=1}^{n} (1 - T_i)} - \frac{\sum_{i=1}^{n} Y_i^B T_i}{\sum_{i=1}^{n} T_i} \right\} \right] / 2 \quad (2)$$

A.2.3 Sources of sensitivity bias

Social desirability bias, also known as sensitivity bias, is formalized by (Blair, Coppock, and Moor 2020). This kind of bias can be seen as a form of measurement error, as the response provided by an individual when they are asked directly about a certain issue or topic is different from the latent true value. Measurement errors can occur for a variety of reasons, such as technical issues, miscommunication between respondents and enumerators, or memory recall mistakes. An additional source of measurement error is generated by the sensitivity of the question. Indeed, respondents may misreport their true opinions or beliefs to avoid embarrassment, to project a favorable image of themselves to others (and to maintain a good self-image), for fear that their responses may be disclosed to authorities, or as a reaction to questions on topics that are considered taboo.

Sensitivity bias can therefore occur if and only if four elements are present: first, a social referent that the respondent has in mind when considering how to respond to a survey question (a social referent could be the respondent themselves); second, a respondent's perception that the social referent can infer the respondent's response to the sensitive question either exactly or approximately; third, a respondent's perception of what response (or nonresponse) the social referent prefers; and fourth, a respondent's perception that failing to provide the response preferred by the social referent would entail costs to themself, other individuals, or other groups. Such costs may be social (e.g., embarrassment), monetary (e.g., fines), or physical (e.g., jail time or personal violence).

These conditions are likely to be met in this specific study as individuals may be embarrassed to disclose to researchers (and to themselves) that they would dislike interacting with gay individuals as employers, co-workers, or customers. The list experiments can address these sources of sensitivity bias by adding an extra level of privacy through non-key items, thus removing the bias by addressing the second element above, that is, making it impossible for researchers to infer the respondent's response. Nevertheless, individuals may still be resistant to answering any questions considered too intrusive or taboo. For this reason, the selected key and non-key items in the main experiments are considered controversial but not so offensive or shocking that the respondents would not consider answering the questions.

A.3 List randomization and list ordering

Both the order of the lists and the order of the items within each list were randomized at the subject level. The order of the statements was randomized at the individual level in both the short and long lists. This randomization served two goals. First, if the order of the items were not randomized and the key statements were listed as last, as done by many papers in this literature, one may worry that seeing a gay-related statement last in three lists could draw extra unwanted attention to the key statements. Second, the order of the statements might also have an impact on the respondents' answers. By randomizing the order, it is possible to eliminate any aggregate effect exerted by the ordering of the statements.

In addition, although it is common practice in the literature not to randomize the order of lists, this survey incorporated some randomization into the design to control for potential order

effects, thus following the approach of (Aksoy, Carpenter, and Sansone 2025). More specifically, the participants were randomly allocated to one of the following six paths:

Path 1: List 1A + KS1, List 1B, List 2A + KS2, List 2B, List 3A + KS, List 3B
Path 2: List 1B + KS1, List 1A, List 2B + KS2, List 2A, List 3B + KS3, List 3A
Path 3: List 2A + KS2, List 2B, List 3A + KS3, List 3B, List 1A + KS1, List 1B
Path 4: List 2B + KS3, List 2A, List 3B + KS3, List 3A, List 1B + KS1, List 1A
Path 5: List 3A + KS3, List 3B, List 1A + KS1, List 1B, List 2A + KS2, List 2B
Path 6: List 3B + KS3, List 3A, List 1B + KS1, List 1A, List 2B + KS2, List 2A

where KS 1, KS 2, and KS3 represent the supervisor, co-worker, and customer key statements, respectively:

"I would feel comfortable supervising a gay employee."
"I would feel comfortable working closely with a gay co-worker."
"I would feel comfortable having a cashier at the supermarket who is gay."

Qualtrics's *Evenly Present Elements* feature was used to randomize the participants into each path.

Lists 1A through List 3B can be seen in the instructions in Online Appendix E. As can be noted from the paths above, half of the participants saw List A first, and the other half saw List B first. When the distributions of answers were compared across these orders via a linear combination test of weighted means (e.g., comparing the responses to each list in Paths 1-2 to those in Paths 3-4 or Paths 5-6, for a total of 36 possible combinations), 9 comparisons presented marginally statistically significant differences, which were all minimal in magnitude (i.e., differences in means of less than 0.2 points). Additionally, when the distributions of answers were compared via unweighted chi-square tests, only 5 out of 36 possible lists presented marginally statistically significant differences.

Notably, while the order of the lists was randomized as just described, the order of the questions in the survey section after the list experiments was the same for all respondents.

A.4 List design

A.4.1 Non-key items in the lists were selected to avoid ceiling and floor effects

While designing the list experiments and choosing the non-key statements, this study followed best practices in the literature (Glynn 2013). For example, it is advised to carefully determine how many non-key statements to include. The number of non-key statements should be neither too low nor too high. The number of key statements should be high enough to avoid a ceiling effect, i.e., participants reporting that all statements are true for them, thus removing the privacy protection provided by the list experiment. At the same time, the number of key statements cannot be too high; otherwise, the respondents may not be able to remember or focus on all statements in the list, thus leading to higher variance and measurement error.

After carefully examining previous studies and noting that (Tsuchiya, Hirai, and Ono 2007) found little impact on list experiment performance when varying the number of non-key items between two and five, this study decided on four non-key statements, as in (Aksoy, Carpenter,

and Sansone 2025). To avoid a ceiling effect, a statement that was expected to be false for most people was included in each list. In addition, each list included a statement that was expected to be true for most people to avoid a floor effect, i.e., participants reporting zero items, thus removing the privacy protection provided by the list experiment.

The remaining two non-key statements were chosen such that they were expected to be negatively correlated: That is, one statement was likely to be supported by more politically conservative people, and another statement was likely to be supported by more politically progressive people. For instance, as shown in Appendix E, the statement "I believe that women should be responsible for the care of children" was expected to be agreed with by conservative respondents, while progressive respondents were expected to support the statement "I believe that it is wrong to apply the death penalty, no matter the crime."

As shown in Figures A1-A3, only a very small share of participants reported the highest and lowest possible items in each of the lists. Thus, it is possible to conclude that the floor and ceiling effects are negligible in these experiments.

Additionally, if the distributions of responses had followed a uniform distribution, then this finding would have indicated that most respondents provided random answers (Coffman, Coffman, and Ericson 2017). It is therefore reassuring to observe that the distributions of responses do not follow such a uniform distribution, as shown in Figures A1-A3.

A.4.2 Negatively and positively correlated non-key items within and between lists

The choice to include negatively correlated items in each list has the additional advantage of decreasing variance and increasing power. High variance is often an issue because the key statement is aggregated with several non-key statements. To some extent, the additional variance is the cost of the higher perceived privacy protection (Glynn 2013). Therefore, list randomization often produces results that are too high in variance to be statistically significant, especially if the attribute, view, or behavior of interest has low prevalence (Karlan and Zinman 2012).

(Osman, Speer, and Weaver 2025) further discuss large standard errors in list experiments and note that because the list randomization method is based on the difference across two variables, the variance of this difference will be mechanically greater than that of a direct question (that is based on only one variable). Indeed, in the absence of any sensitivity bias, the standard error on the direct question estimate of the proportion of respondents not feeling comfortable with gay individuals would be the following:

$$\sqrt{\frac{p(1-p)}{N}}$$

where p is the proportion answering yes and N is the sample size.

Holding the sample size constant and assuming that the proportion of people answering yes to the *k* non-key questions was also p and that the answers were independent, the standard error on the list randomization would be the following:

$$\sqrt{\frac{(2k+1)p(1-p)}{N}}$$

For this reason, the standard error from the list randomization is mechanically larger than that from the direct question. The variance expressions for both the direct question and the list experiment estimator in the presence of sensitivity bias and/or unbalanced design are instead derived by (Blair, Coppock, and Moor 2020).

To further increase power in the double list experiment, the non-key statements in Lists A and B were also designed to be positively correlated across lists (Glynn 2013). For example, the statement "I believe that women should be responsible for the care of children" in List 1A was chosen to be positively correlated with the statement "I believe that the poor make little effort to get out of poverty" in List 1B (as reported in Appendix E).

A.4.3 Statistical Power Calculations

The statistical power analysis for the double list experiment builds on the conventional two-sample independent groups framework but adapts the calculations to a paired design where each of the 4,000 respondents provides both a control and a treatment measurement. In the independent groups design, that is, the standard between-subject control and treatment design, the minimum detectable effect size (MDE – Cohen's d) is given by the following:

$$d = \left(z_{\left(1-\frac{\alpha}{2}\right)} + z_{(1-\beta)}\right) \times \sqrt{\frac{2}{n}}$$

For a two-tailed test with α = 0.05 ($z_{\left(1-\frac{\alpha}{2}\right)}$= 1.96) and 80-percent power ($z_{(1-\beta)}$ = 0.84) and with 2,000 respondents per group (n = 2,000), this yields the following:

$$d = (1.96 + 0.84) \times \sqrt{\frac{2}{2000}}$$

$$= 2.80 \times 0.03162$$

$$\approx 0.0885$$

However, in a double list experiment, every respondent contributes data in both conditions—effectively serving as control and treatment observations—which significantly improves precision. For a paired design, the minimum distance estimator (MDE) formula is modified to account for the within-subject correlation (ρ) between the control and treatment measurements:

$$d = \left(z_{\left(1-\frac{\alpha}{2}\right)} + z_{(1-\beta)}\right) \times \sqrt{\frac{2(1-\rho)}{n}}$$

Here, N is the total sample size (4,000), and ρ quantifies the degree to which the responses from the two lists are similar for each respondent. A higher ρ indicates that much of the

individual variation is common to both conditions, reducing the effective variance of the difference. For example, assuming a moderate correlation of $\rho = 0.5$, one obtains the following:

$$\sqrt{\frac{2(1-0.5)}{4000}} = \sqrt{\frac{1}{4000}} \approx 1/\sqrt{4000} \approx 0.01581$$

$$\text{Thus, } d = 2.80 \times 0.01581 \approx 0.0443.$$

Even if one assumes no correlation ($\rho = 0$), the standard error becomes $\sqrt{\frac{2}{4000}} \approx 0.02236$, leading to $d \approx 0.0626$. Therefore, this design is capable of detecting effect sizes in the range of approximately d = 0.044 to d = 0.063.

The introduction of the correlation parameter is critical in paired designs because it directly influences the variance of the difference between the two measurements. When responses are highly correlated, the variability that is attributable to individual differences is largely canceled out, thereby increasing statistical power to detect subtle effects (Cohen, 1988; Vickers, 2001). By incorporating ρ into the power analysis, one can more accurately reflect the efficiency gains from the within-subject design of the double list experiment, in addition to the gains resulting from observing the treated outcome for all respondents (i.e., as opposed to within-subject designs). This ensures that even modest latent differences in discriminatory attitudes can be detected with this study's robust sample size.

A.4.4 Sensitive versus non-sensitive non-key items

In line with (Chuang et al. 2021), to draw less attention to the key statements and increase the validity of the list experiment, some of the non-key statements in the lists were political or sensitive in nature. For example, as reported in Appendix E, the lists included items about gender norms, immigration (refugees), the death penalty, abortion, taxes, drugs, environmental protection, law enforcement, social protests, and sex education in schools.

Additionally, in line with (Berinsky 2004), the participants were not provided a "don't know" option in the direct question since individuals who held socially stigmatized opinions may have hidden their opinions behind a "don't know" response.

Finally, (Coffman, Coffman, and Ericson 2017) showed that list experiments work better when the stigmatized answer in the related direct question is a "no" instead of a "yes." Thus, the key items and related direct questions were designed such that the socially stigmatized answer was always a "no."

A.4.5 List example provided to the participants

To ensure that the respondents understood the task, they were presented with an example of a list experiment question and guided through the reasoning behind a certain numerical answer. They were reminded that the question asked them to report only the number of items that were true for them, not which ones.

All respondents were then given the option to review the instructions one more time before answering the six list experiment questions.

A.4.6 Priming

Both the recruitment messages (reported in Appendix E.1) and the consent forms (reported in Appendix E.2) stated that the goal of the survey was to understand individuals views and preferences. The description of the study did not specifically mention LGBTQ+ issues to avoid priming the respondents or obtaining a self-selected sample. A total of 207 participants (5.17 percent of the sample) clicked on the link to read the entire consent form.

A.4.7 Placebo tests in the previous literature

Importantly, increased reporting under the veil of the list experiment is not simply mechanical. Indeed, previous research has shown that list experiments provide increased estimates of prevalence only for stigmatized views: there is no evidence of this technique leading to an increase in reporting of innocuous behaviors (Tsuchiya, Hirai, and Ono 2007; Coffman, Coffman, and Ericson 2017). For instance, (Coffman, Coffman, and Ericson 2017) did not find any significant misreporting when the additional key statement in the longer list was "It has rained once where I live in the last four days."

A.4.8 Advantages of survey experiments and online surveys

As emphasized by (Stantcheva 2023), not only are surveys a way of collecting data, but they also allow researchers to create their own identifying and controlled variation, thus providing a high level of control in the data generation process. In addition, while administrative data are excellent resources, it must also be acknowledged that, unlike surveys, these data cannot capture factors such as perceptions, beliefs, attitudes, knowledge, or reasoning. Similarly, while economists often favor the revealed preference approach, many crucial determinants of social, economic, and political outcomes – such as perceptions, beliefs, attitudes, and reasoning – are not always easily inferred from observed behavior. This limitation makes surveys a valuable complement to both administrative and other observational data. This is especially relevant in the context of discrimination, where identifying its sources solely through aggregate statistics and observational data can be challenging (Domínguez, Grau, and Vergara 2022).

(Stantcheva 2023) also specifically highlighted some advantages of online surveys, as opposed to in-person, phone, or mail surveys, in terms of selection. First, online surveys give people the flexibility to complete the survey at their convenience, which reduces selection based on who is free to answer during regular work hours or who opens the door or picks up the phone. This feature may allow individuals who need to juggle different responsibilities (e.g., carers) to take part in a study. Second, the convenience of mobile technologies may entice some people who would otherwise not want to complete questionnaires or answer questions on the phone to take part in surveys. Third, online surveys can reach people who would be hard to interview in person (e.g., younger respondents, those who often move residences, respondents in remote or rural areas). Fourth, platforms that administer this kind of survey offer a variety of rewards for taking part in surveys, which can appeal to a broader group of people.

A.5 List experiment assumptions

The validity of a list experiment relies on three assumptions: 1) treatment randomization, 2) no design effect, and 3) "no liar." The first assumption means that the sample is split at random. The second assumption means that respondents do not give different answers to non-key statements depending on whether they are in the long list group. The third assumption means that respondents answer the key statement truthfully.

A common practice to check the first assumption – treatment randomization – is to test for differences between the short list and long list groups' responses to important variables in the survey. More precisely, since this study is based on double list experiments, one must check whether the participants treated in List A are systematically different from those treated in List B. Table A1 checks the differences between the two groups in terms of their demographic covariates. In general, there are no significant differences between the two groups, with the exceptions of for slight differences (marginally statistically significant and negligible in magnitude) for the 50-64 age group and for individuals making more than approximately USD 1,500. This evidence is thus reassuring that the randomization of treatment was effective.

Moreover, in line with (Gerber and Green 2012) and (Detkova, Tkachenko, and Yakovlev 2021), the main analysis does not rely on only mean comparisons; also instead, it employs regression analyses that control for observable characteristics (as discussed in Section 3.1), including participant age, sex, sexual orientation, and gender identity, as well as indicators for ethnicity (i.e., African descent) and indigenous status, region and commune (for those in the Metropolitan Region of Santiago) fixed effects, educational level, income, employment status, managerial experience, religious views, and political affiliation. Additional controls include whether at least one child less than 18 years of age lives in the participant's household, the number of people living in the participant's household, marital status, LG comfort (as defined in Appendix B), an indicator for whether the participant knows a gay person (as defined in Appendix B), and day-of-week and week-of-sample indicators.

The second assumption – no design effect – requires the respondents not to change their answers to non-key statements depending on whether the key statement appeared in the list (i.e., whether they saw the long list). To clarify, suppose that a respondent in the short list group answered two non-key statements affirmatively. If the respondent had been assigned to the long list group, their answer must have been either "2" or "3" (that is, the respondent either answered two non-key statements affirmatively or answered two non-key statements plus the key statement affirmatively). Notably, it is not assumed that the respondents gave truthful answers to these non-key statements; it is only assumed that the answers were consistent in the short and long list groups. (Blair and Imai 2012) proposed a statistical test for the no design effect assumption, and this test can be implemented via the Stata command *kict deff* (Tsai 2019). The first step is to estimate the probabilities of all possible types of item count responses. If some of these estimated probabilities were a nonsensical value (e.g., a negative value), this result would raise doubts about the validity of the no design effect assumption. One can then test whether such negative estimates have arisen by chance.

In line with (Blair and Imai 2012) and (Tasi 2019), the no design effect tests were performed twice, with and without the use of the method of generalized moment selection (GMS). Using GMS, 6 (out of 60) estimated probabilities of the item count responses present negative values, and 2 (out of 12) Bonferroni-adjusted p-values are slightly above conventional significance thresholds. These marginally significant p-values correspond to List B (supervisor) and List A (cashier). Conducting the same tests without the use of the GMS method results in the same number of estimated probabilities of the item count responses presenting negative values, but all Bonferroni-adjusted p-values are above conventional significance thresholds. That is, all tests fail to reject the null hypothesis of no design effects. Hence, one cannot reject the null hypothesis that such negative estimates have arisen by chance. Overall, it is possible to conclude that the available evidence broadly supports the no design effect assumption.

It is not statistically feasible to check the "no liar" assumption, not only because the respondents' answers to the key statement are, by design, unobserved but also because their truthful answers are unknown (otherwise, there would be no point in using the list experiment technique). This study tried to limit any concerns about this assumption by running these experiments on an online anonymized platform and by ensuring that, when designing the lists, agreeing to all or none of the statements would be highly unlikely. Indeed, Figures A1-A3 present the distribution of responses for each list and key statement: the modal response in most lists is 2. Moreover, as noted in Appendix A.4.1, the percentage of times where the responses are 0 or 4 (5 for long lists) is negligible, meaning that the privacy of responses was protected.

A.6. Data quality checks

A.6.1 Procrastinators and speeders

As done by (Gutiérrez and Rubli 2024), it is possible to check the robustness of the main list experiment findings by excluding participants who completed the study very quickly or very slowly since they may not be paying as much attention to the study instructions. The median respondent took 848 seconds (14.1 minutes) to complete the list experiments. The results presented in Figure A4 show that the main findings are robust to the exclusion of 401 participants who took less than 491 seconds (8.1 minutes, top 5 percent) or more than 3,354.5 seconds (55.9 minutes, bottom 5 percent) to complete the study.

A.6.2 Attention and quality checks

The list experiments included an attention check, and the survey included two additional attention check questions. In line with the recommendation of (Haaland, Roth, and Wohlfart 2023), at the beginning of the study, the rationale for including such attention checks was explained to the respondents. This explanation can mitigate concerns about the participants' negative emotional reactions to the use of attention checks. More specifically, the respondents were informed that, sometimes, there are participants who do not carefully read the questions and just quickly click through surveys, thus resulting in random answers that compromise the research findings. For this reason, they were told that the study included several attention checks and that failing to complete two or more of these questions correctly may cause them to be ineligible for compensation.

A total of 776 participants (19.4 percent of the main sample) failed one out of the three attention checks. In particular, 761 of these 776 participants failed the first attention check, which was arguably the most challenging check. A very low share of participants, 0.88 percent (35 respondents) failed two or more attention checks. For robustness, Figure A5 presents the results when participants who failed at least one attention check are excluded. The main results are robust to the exclusion of these participants.

It is also possible to check whether some respondents provided the same answer to all list experiment questions (which might be an indication of inattention). Across all seven lists, 12 participants (0.3 percent of the main sample) provided the same number. In the first six lists (thus excluding the list that serves as an attention check), 58 participants (1.45 percent) provided the same number for all four lists. As shown in Figure A6, the main findings are robust to the exclusion of these 58 participants.

Furthermore, as suggested by (Stantcheva 2023), all participants were required to complete a Captcha (Completely Automated Public Turing Test to tell Computers and Humans Apart) test to verify that they were humans. Figure A7 tests the robustness of the main findings to the exclusion of participants who scored below 0.5 in the Captcha Verification test. It is generally accepted that scores below 0.5 suggest a higher risk that a respondent is not human, with 0.0 posing the highest risk and 1.0 posing the lowest risk. Out of the sample of 4,000 respondents, 3,850 scored 0.5 or higher. The results are robust to the exclusion of the participants who scored less than 0.5.

In addition, the Qualtrics software collected basic geographic information to verify that all respondents who completed the survey were indeed from Chile. Figure A8 presents the main findings excluding 98 participants who completed the study from outside of Chile according to the Qualtrics approximation of the respondents' latitude and longitude and 62 participants for whom Qualtrics did not provide a location approximation. The results are robust to the exclusion of these responses.

A.6.3 Attrition rate

As stated in the consent form, the participants were always free to leave the study whenever they wished. The attrition rate was 23.98 percent; that is, 1,251 participants started the study but did not complete it. This attrition occurs overwhelmingly *before* the start of the list experiment. Specifically, 775 participants (61.95 percent of all dropouts) exited the study before the start of the list experiment (i.e., during consent, overview, instructions or the example list experiment questions). Out of the remaining participants, 144 (11.5 percent) dropped out *during* the list experiment. The remaining respondents who proceeded to complete the survey did not seem to display a pattern of abandoning the study more frequently after specific survey questions. Figure A9 plots the number of non-missing values for variables at specific points through the study. This figure illustrates the patterns of attrition mentioned above, that is, that dropouts occurred mostly *before* the start of the list experiment. Furthermore, without considering individuals who did not start the list experiments, the attrition rate was only 10.6 percent, which is much lower than that in surveys conducted by national statistical offices.

Since the overwhelming majority of dropouts did not reach the demographic portion of the survey, it is only possible to compare the demographic characteristics between the main sample (unweighted statistics for comparison) and the small subset of dropouts who reached the latter parts of the study. Performing this limited comparison, this study finds that there are several similarities across the samples. For example, 47 percent of respondents were female in the main sample, while 46.9 percent (68 participants) were female among the 145 dropouts who responded to the sex question. This difference is not statistically significant. Similarly, 76.9 percent of the main sample respondents reported having managerial experience. The same number is 78.64 percent for the sample of dropouts who reached this question (220 participants). This difference is also not statistically significant. In the main data, 36.05 percent of respondents were not from the Metropolitan Region of Santiago, while the same statistic for the sample of dropouts is 41.03 percent. This difference is also not statistically significant.

However, there are some differences across age, employment, and income. While the mean and median age in the main sample is 47.6 and 46 years, respectively, the limited sample of dropouts who reached the age question (269 participants) had a slightly higher mean and median age, 52.7 and 52 years, respectively. This difference is statistically significant at the 99-percent level. Similarly, 84.95 percent of the main sample reported being employed, while only 75 percent (165 out of 220) reported being employed in the sample of dropouts. This difference is statistically significant at the 99-percent level. Finally, in the main sample, 61.2 percent of participants reported an income higher than approximately USD 1,500, while 53.17 percent (67 participants) of dropouts who answered the income question (126 participants) reported the same income levels. This difference is statistically significant only at the 90-percent level.

By using the sample of eventual dropouts who managed to complete the list experiment (290 respondents), it is possible to estimate the level of support for the three key statements among these participants. For the supervisor key statement, the double list experiment estimate is 44.62 percent. For the co-worker key statement, the double list experiment estimate is 46.16 percent. Finally, for the cashier key statement, the double list experiment estimate is 48.30 percent. Given that the main estimates via unweighted data (Figure A11) are 62.7 percent, 66.8 percent, and 72.9 percent, respectively, the findings indicate that these dropouts may hold more conservative views regarding sexual minority men. However, including the sample of dropouts who completed the list experiment in the main estimates (Figure A10) provides results that are remarkably similar to those in Figure A11. Overall, these results, paired with the fact that attrition overwhelmingly occurred before the list experiment, substantially decrease any concerns that participants may have dropped out of the study systematically.

The main sample uses data only from the participants who completed the entire study. The main sample of 4,000 respondents does not include any of the 1,251 participants described above who dropped out at various points before finishing the study.

A.6.4 Feedback questions

At the end of the survey, the participants were asked to report the extent to which they agreed with the statement "The instructions were clear." The menu of answers consisted of a 5-point Likert scale ranging from "totally agree" to "totally disagree." As shown in Table A2, over 97 percent of participants indicated that they totally or somewhat agreed with the statement, suggesting that most respondents found the instructions to be clear. However, for robustness, Figure A12 shows the main results excluding 63 participants who responded "neither agree nor disagree," 25 participants who responded "somewhat disagree," and 16 participants who responded "totally disagree" with the statement "The instructions were clear." The results remain robust to the exclusion of these 104 respondents.

This 5-point Likert scale question was followed by two open-ended questions to provide the respondents with a venue to leave feedback about the study to the researchers. The first open-ended question asked the respondents to report anything that may have been unclear or that could have been confusing during the study. A total of 2,385 participants provided responses to the prompt "Is there anything that is unclear or confusing in the study?" The responses were first processed by cleaning the text (converting to lowercase, removing punctuation, and filtering out stopwords) and then categorized into thematic groups. This analysis reveals that over half of the respondents (56.23 percent) indicated "no issues," suggesting that the study was largely perceived as clear. Additionally, 205 participants (8.60 percent) offered explicit "Positive Feedback" about the study's clarity. A small fraction (1.38 percent) raised "clarity issues," while 0.80 percent provided "suggestions" for improvement, and 13.50 percent of responses fell into the "mixed" category, reflecting overlapping or ambiguous feedback. Finally, 19.50 percent of the responses were "uncategorized," as they did not neatly align with any of the predefined themes. Among those who raised issues, their chief complaint was about the purpose of the study: An additional manual inspection of a randomly selected subset of 100 respondents indicates that several participants were confused about the objective of the study and the reason why the list experiment questions were asked. These comments are in line with the study objective not to prime the respondents by disclosing the aim of the survey.

The second and last open-ended question asked the participants whether there was anything else that they wanted to share with the investigators. A similar analysis found that, out of 2,307 non-missing responses to the prompt "Is there anything else that you would like to share with the investigators?," the majority (53.75 percent) simply provided "no comments," indicating that they had nothing further to share. A substantial percentage of responses (37.62 percent) was "uncategorized," meaning that they did not neatly fit into the predefined thematic patterns. In contrast, smaller proportions conveyed distinct sentiments: 1.73 percent were flagged as "negative," 3.12 percent were categorized as "positive," 2.60 percent offered "suggestions" for improvement, and 1.17 percent were classified as "mixed." Most respondents offered minimal additional feedback in response to this question, and inspections of a randomly selected subset of 100 responses uncovered the diversity of comments, ranging from questions about specific portions of the survey to personal anecdotes related to the content of this study.

Overall, the analysis of these final survey questions indicates no significant concerns regarding the survey design or data quality

A.6.5 Standard errors and weights

Standard errors were computed in line with (Glynn 2013): because estimation is conducted by taking the difference in mean responses between two independent sets of respondents, the variance of the estimator can be calculated with the standard large-sample formula for a difference in means, and confidence intervals can be computed in the usual fashion.

In addition, the main estimates and standard errors reported in Table D1 do not change substantially when the Stata command *kict ls* (Tsai 2019) is used to perform least squares estimation specifically for a double list experiment.

The main sample is weighted based on region, sex, age and educational level. The local partner, Datavoz, conducted the weighting via parameters from population projection data for 2024 (for the region, sex, and age variables) and from the 2022 census data (for the education variable). Datavoz used the *raking*, also known as random iterative method (RIM) weighting, method, which consists of iteratively adjusting the proportional weights for each observation until the sample distribution matches the population distribution for the variables of interest. Before weighting, the raw data oversampled individuals who, compared to the general Chilean population, attained higher levels of education, respondents over 34 years of age, and residents of the Metropolitan Region of Santiago. Following guidance from Datavoz, *pweights* were used when the Stata command allowed it, otherwise *aweights* were used.

Figure A11 further shows that the main findings are robust to the use of unweighted data, although the estimated share of people who are comfortable with a gay cashier from the list experiments is higher in the unweighted data.

A.6.7 Datavoz Pilot and Soft Launch

After the experiment was piloted on Prolific with a sample of Chilean participants (discussed in detail in Appendix C), the local partner, Datavoz, conducted a pilot and soft launch of the study with the same target population as the main sample before launching the main data collection effort. The Datavoz pilot consisted of 235 responses collected in September 2024, and the soft launch consisted of 62 responses collected in October 2024. The purpose of these exercises was to test the survey instrument in the field, gather initial feedback from participants, and adjust the questionnaire and Qualtrics code as needed.

As outlined in the pre-analysis plan, Figure A13 shows the main list results with the inclusion of data from both the pilot and soft launch, amounting to a total of 4,297 observations. The unweighted results including these additional samples largely mimic those presented for the unweighted main sample in Figure A11.

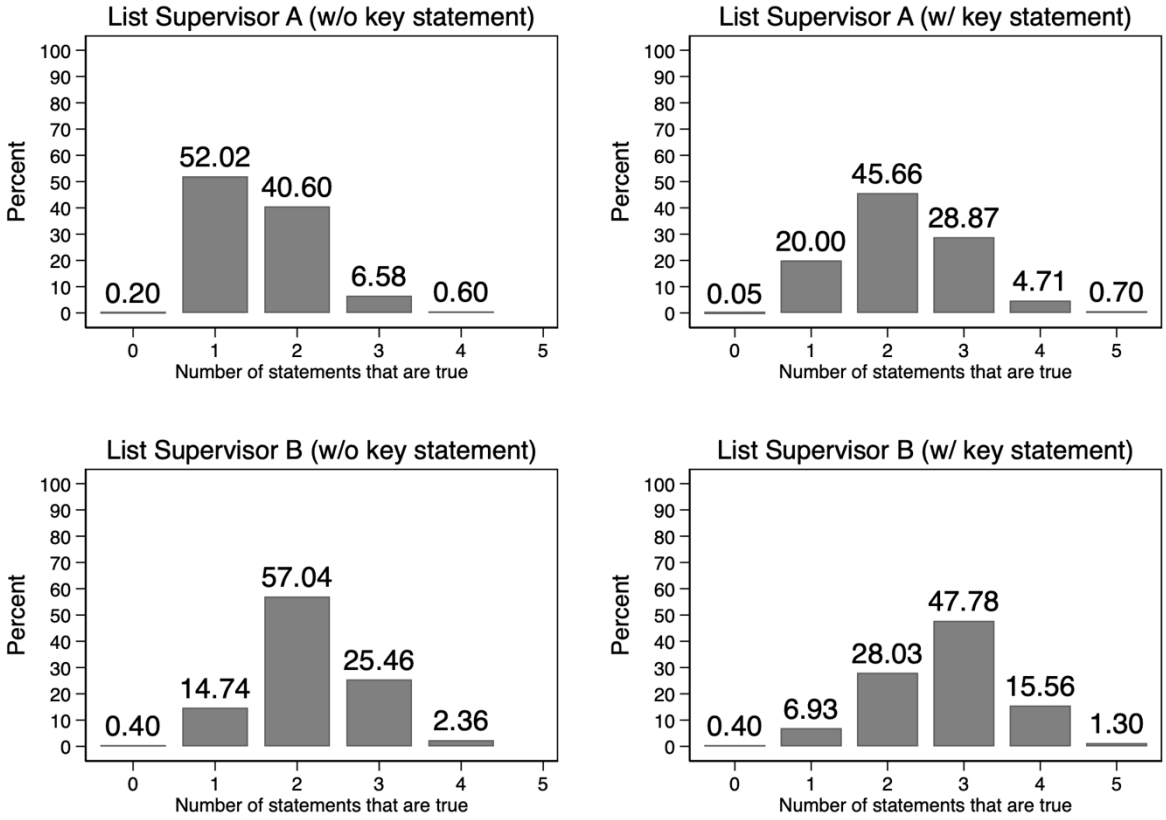
A.7.7 Additional considerations

Before pre-registering the experiment and starting the data collection, the pre-analysis plan and survey were reviewed by four experts on list experiments. The questionnaire and experiment were then modified to incorporate their recommendations.

Finally, this study relies on list experiments instead of the randomized response technique (where respondents use a private randomization device – e.g., flipping a coin – to determine

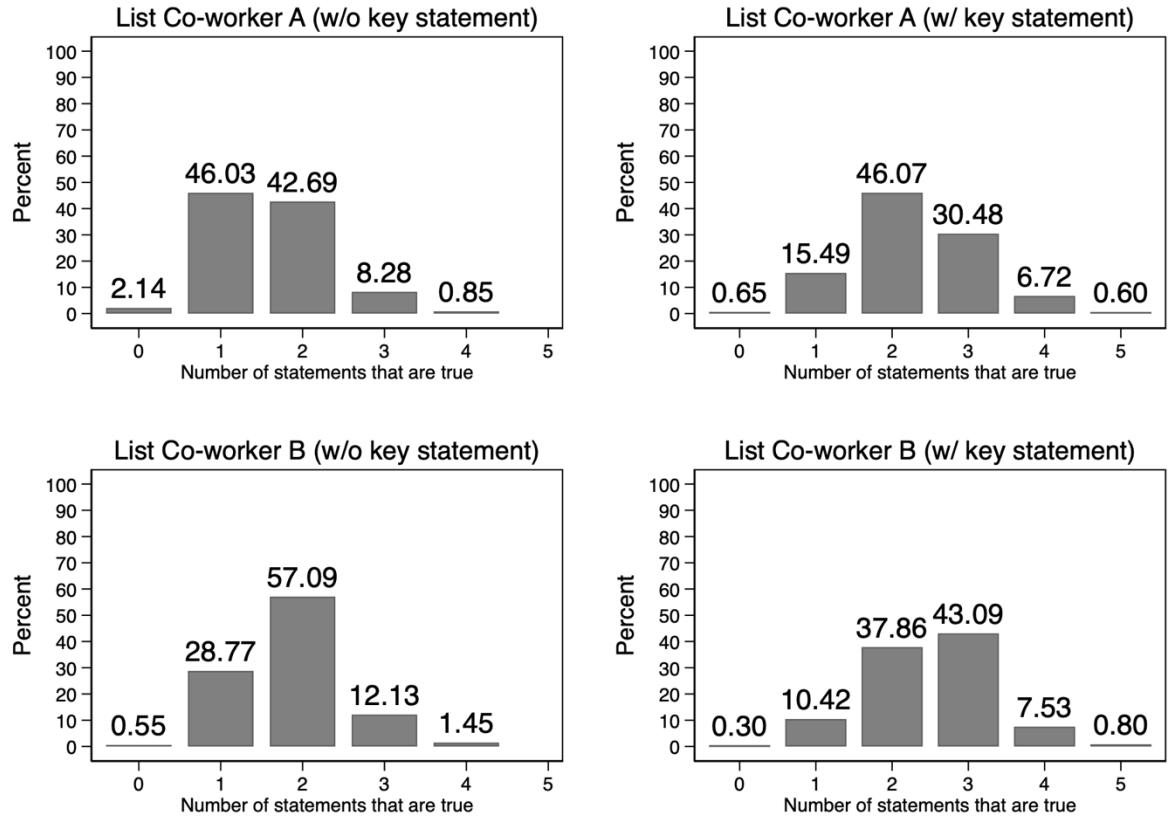
whether they answer either a sensitive question or an innocuous question) for three reasons in addition to the advantages of the list experiment methods discussed in Section A.4. First, the randomized response technique is more difficult to implement online. Second, research suggests that participants trust the randomized response technique less than they trust list experiments (Coutts and Jann 2011). Third, research also documents that participants may not respond to the randomization device relied upon by this technique as instructed (John et al. 2018).

Figure A1. Responses to the Supervisor Lists with and without the Supervisor Key Statement



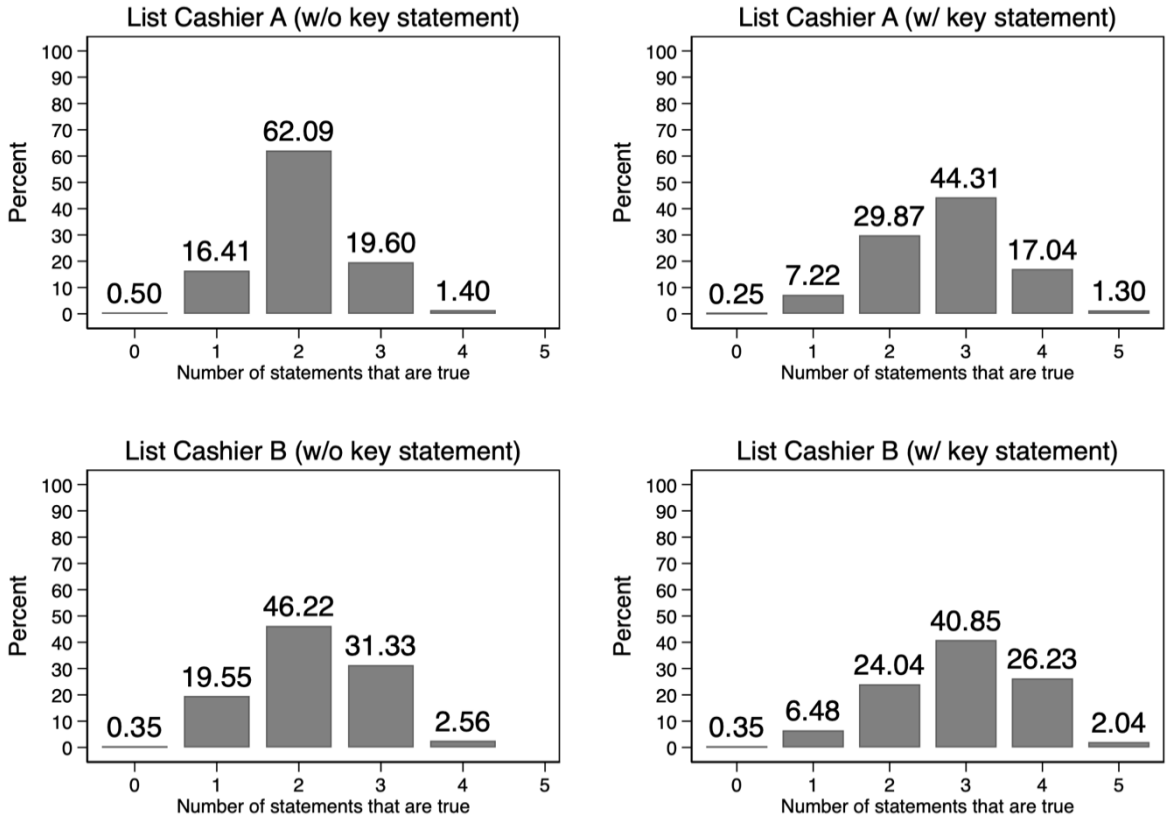
Notes: Unweighted statistics. Key statement: "I would feel comfortable supervising a gay employee." Number of observations: 2,005 (List A without the key statement), 1,995 (List A with the key statement), 1,995 (List B without the key statement), and 2,005 (List B with the key statement).

Figure A2. Responses to the Co-worker Lists with and without the Co-worker Key Statement



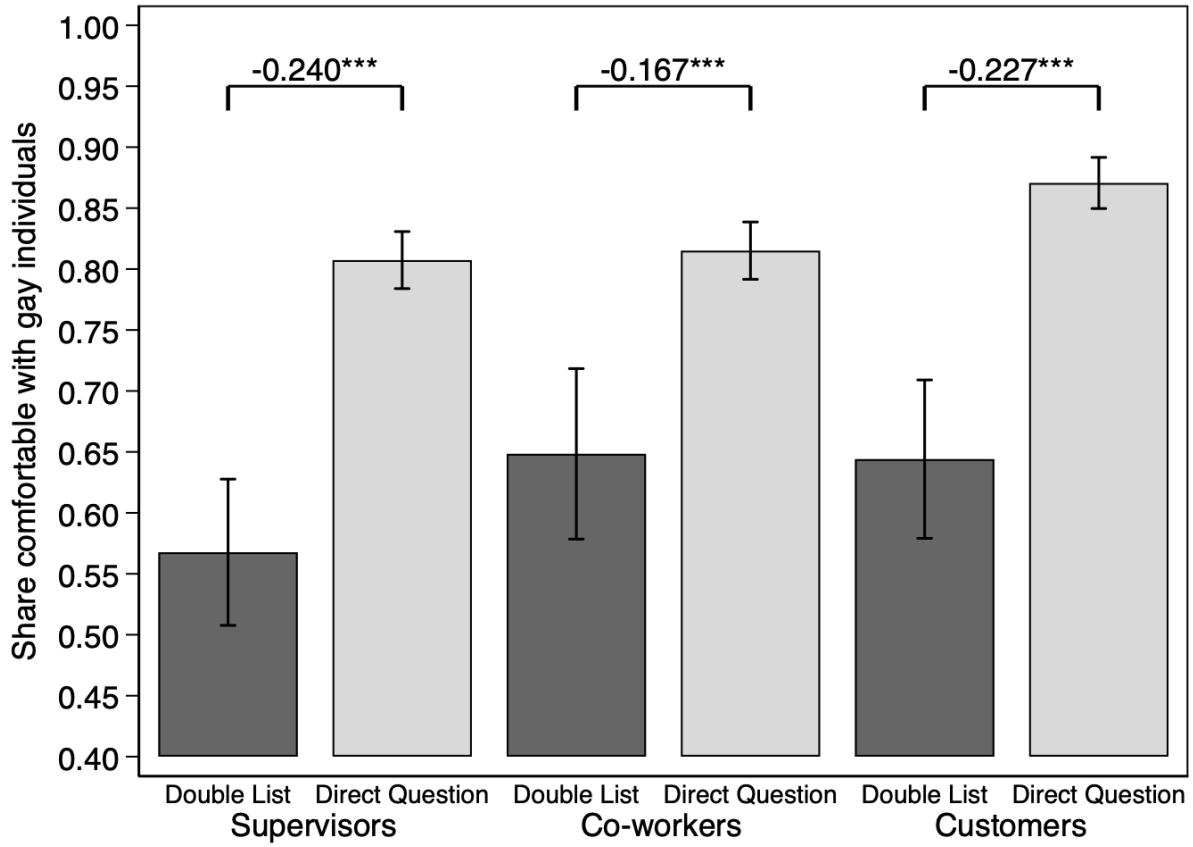
Notes: Unweighted statistics. Key statement: "I would feel comfortable working closely with a gay co-worker." Number of observations: 2,005 (List A without the key statement), 1,995 (List A with the key statement), 1,995 (List B without the key statement), and 2,005 (List B with the key statement).

Figure A3. Responses to the Cashier Lists with and without the Customer Key Statement



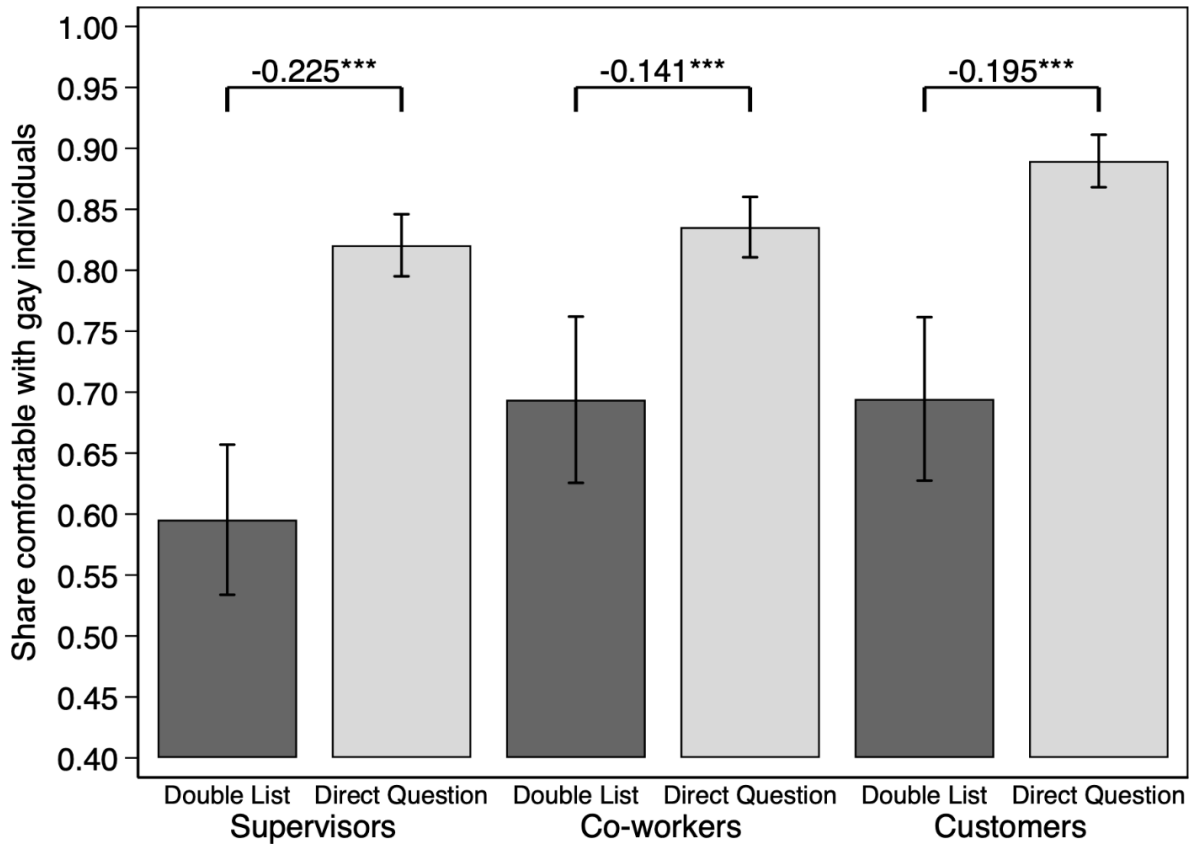
Notes: Unweighted statistics. Key statement: "I would feel comfortable having a cashier at the supermarket who is gay." Number of observations: 2,005 (List A without the key statement), 1,995 (List A with the key statement), 1,995 (List B without the key statement), and 2,005 (List B with the key statement).

Figure A4. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Speeders and Procrastinators



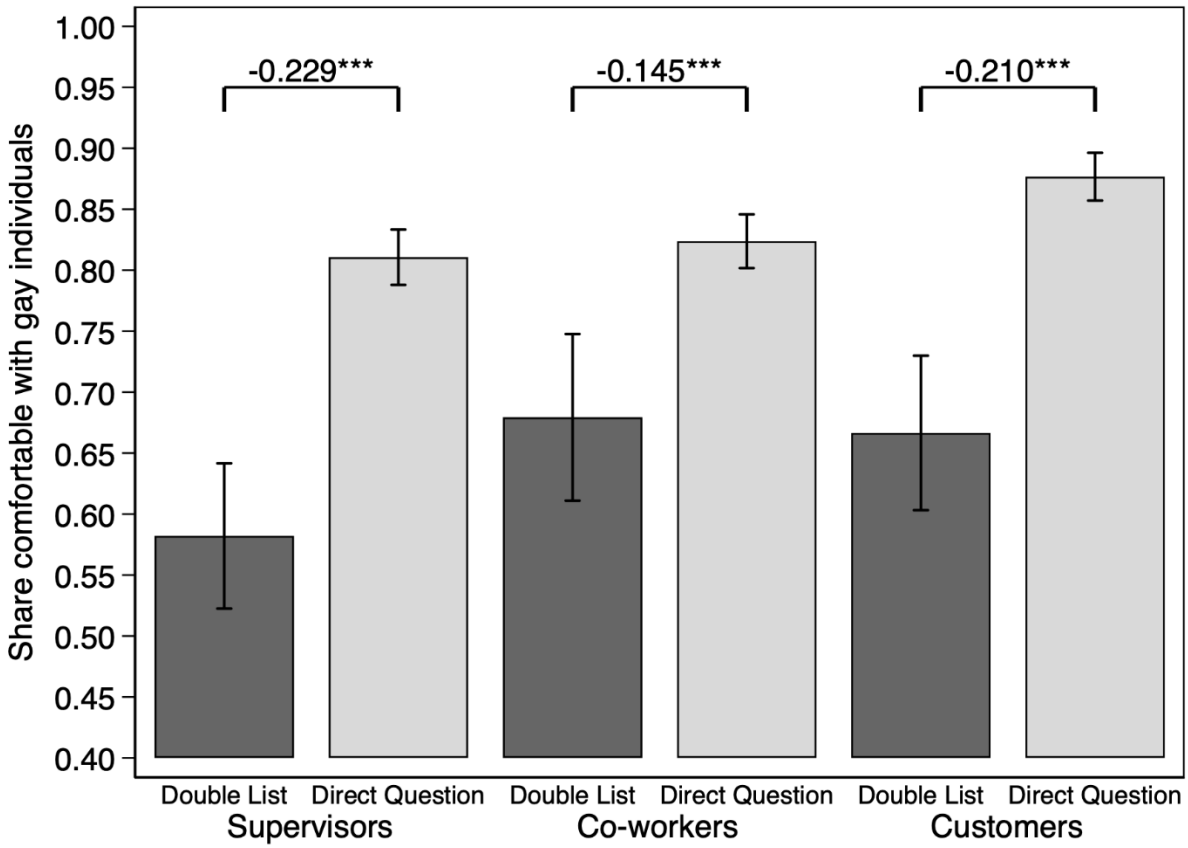
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,599. The sample does not include 401 participants who took less than 491 seconds (8.1 minutes, top 5 percent) or more than 3,354.5 seconds (55.9 minutes, bottom 5 percent) to complete the study. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A5. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Participants Who Failed One or More Attention Checks



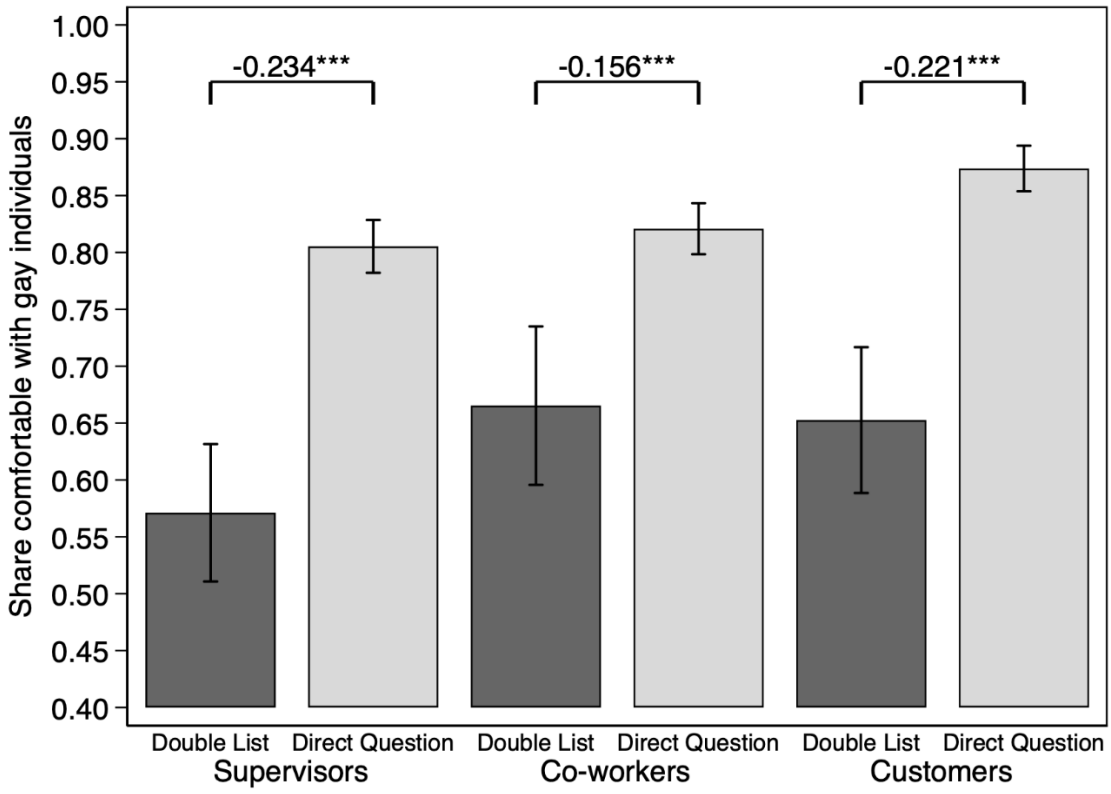
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,224. The sample does not include 776 participants who failed at least one attention check. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A6. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Participants Who Provided the Same Answer in All Six Lists



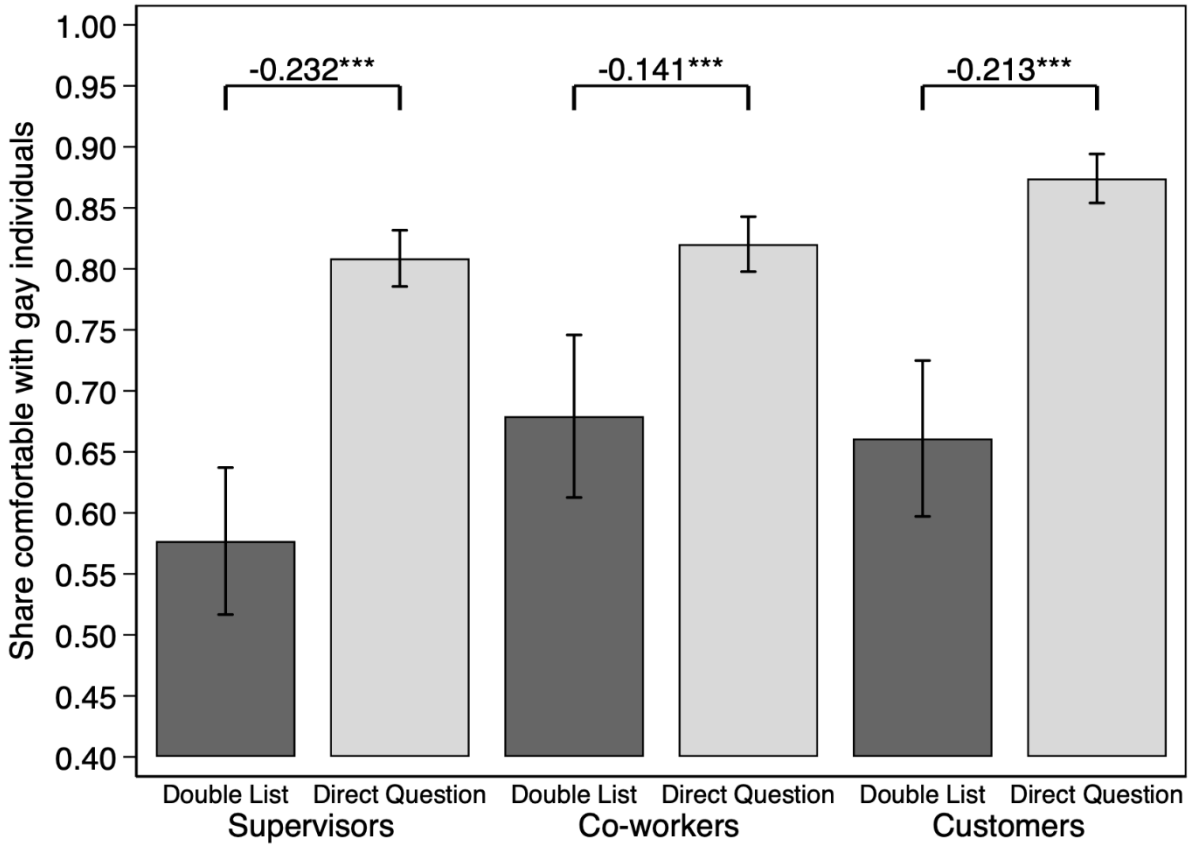
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,942. The sample does not include 58 participants who provided the same response in all six lists. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A7. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Participants Who Scored Less Than 0.5 in the Captcha Verification



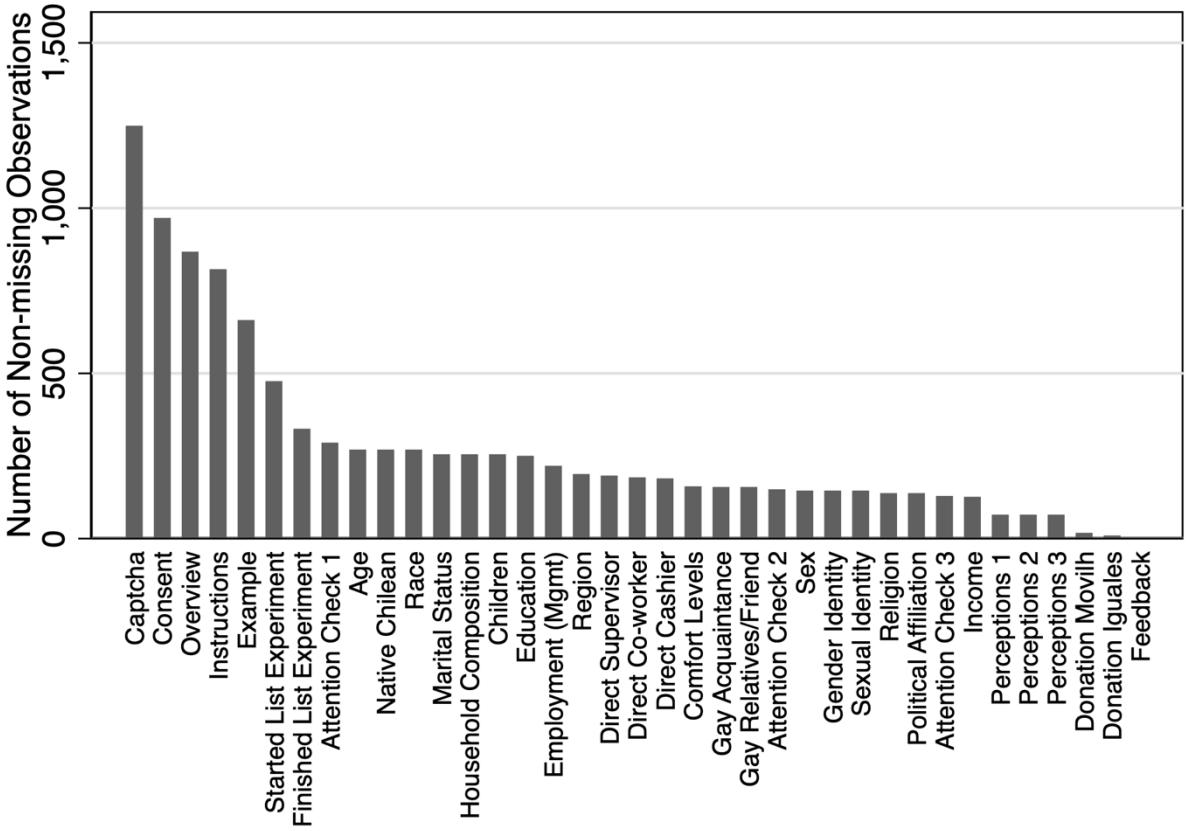
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. The sample does not include 150 participants who scored less than 0.5 in the Qualtrics Captcha (Completely Automated Public Turing Test to tell Computers and Humans Apart) verification at the beginning of the study. In Qualtrics, Captcha is a third-party service provided by Google. Number of observations: 3,850. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A8. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Participants Who Responded Outside of Chile or for Whom There Were No Location Data



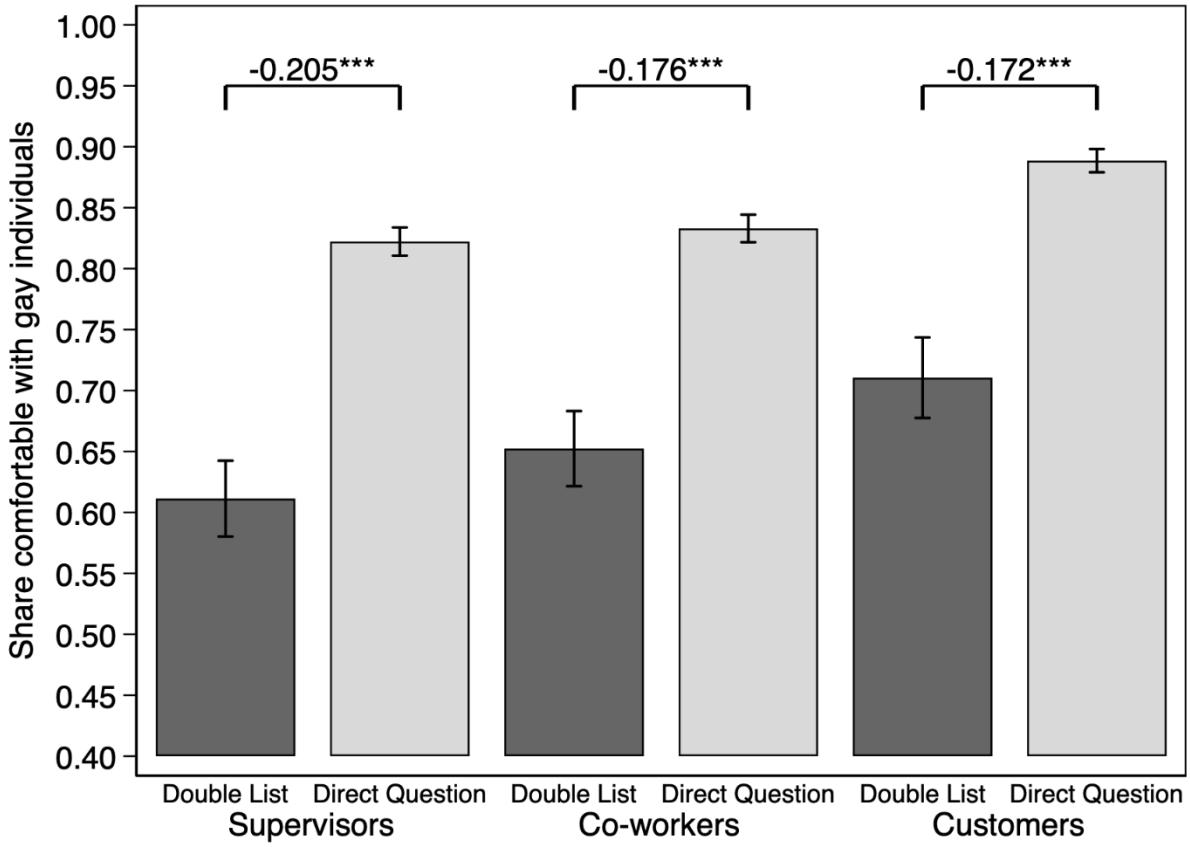
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,840. The sample does not include 98 participants who responded to the survey from outside of Chile and 62 participants for whom Qualtrics did not capture location data. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A9. Attrition Numbers by Survey Question



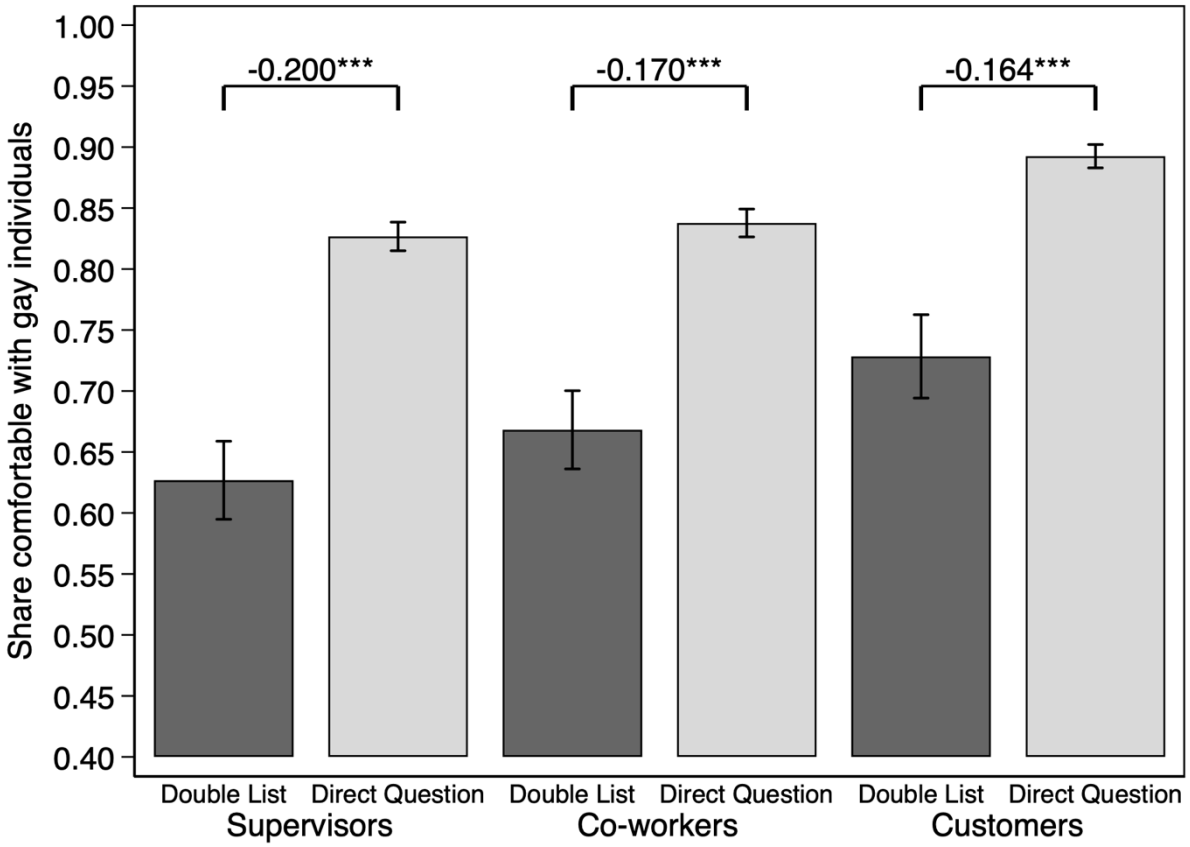
Notes: Unweighted statistics. Number of starting observations: 1,251. The variables listed in the horizontal axis appear sequentially in the order in which they were presented in the study.

Figure A10. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Including Participants Who Finished the List Experiments and Later Dropped Out of Study



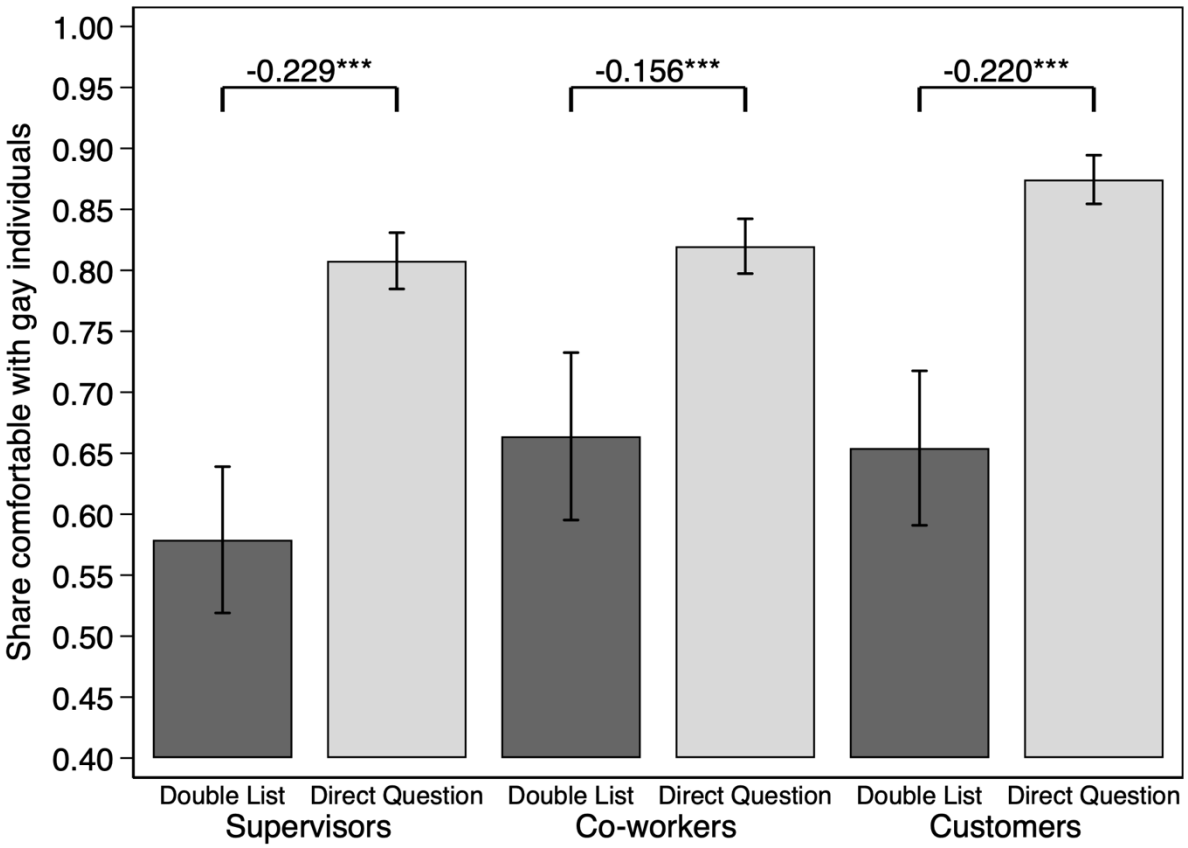
Notes: Unweighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations for direct questions: 4,190 (supervisors), 4,185 (co-workers), and 4,182 (customers). The sample includes participants who were excluded from the main study because they dropped out of the survey. The number of observations varies across supervisors, co-workers, and customers because not all dropouts who finished the list experiment reached the direct questions. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A11. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Unweighted Data



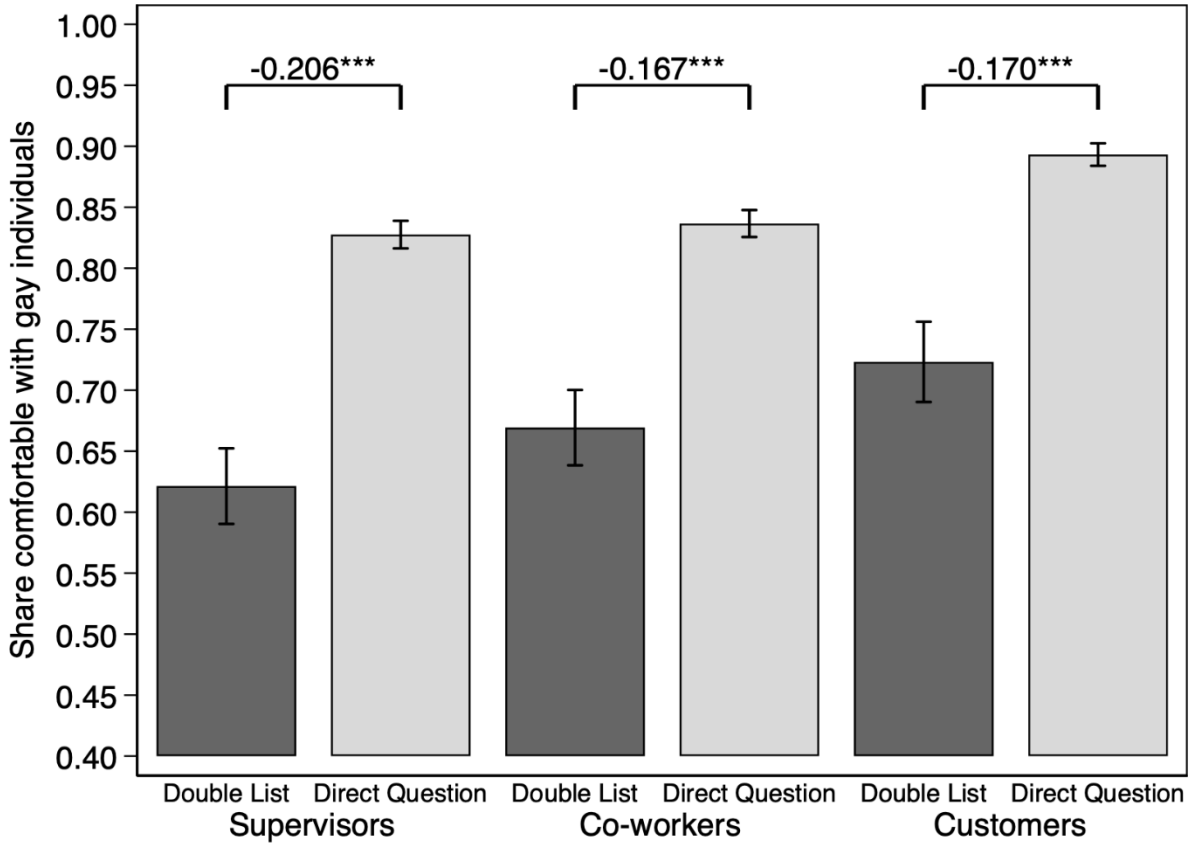
Notes: Unweighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 4,000. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A12. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Excluding Participants Who Did Not Agree with "The Instructions Were Clear" Statement



Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. The sample does not include 63 participants who responded "neither agree nor disagree," 25 participants who responded "somewhat disagree," and 16 participants who responded "totally disagree" with the statement "The instructions were clear." Number of observations: 3,896. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure A13. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Including Data from the Datavoz Pilot and Soft Launch



Notes: Unweighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. The sample includes data from a first pilot (N=235) conducted in September 2024, a soft launch (N=62) conducted in October 2024, and the main wave (N=4,000) conducted from November to December 2024. Total number of observations: 4,297. *p < 0.10; **p < 0.05; ***p < 0.01.

Table A1. Balance Table

Variable	Treatment A Mean	Treatment B Mean	Difference	p-value
Age	47.61	47.77	0.16	0.678
Between 18 and 34	0.131	0.125	-0.006	0.594
Between 35 and 49	0.477	0.457	-0.019	0.221
Between 50 and 64	0.280	0.313	0.033	0.022**
65 or older	0.113	0.105	-0.008	0.414
Sex: Female	0.477	0.463	-0.014	0.380
African Descent	0.032	0.027	-0.004	0.438
Indigenous	0.067	0.066	-0.001	0.916
Gender Identity and Sexual Orientation				
Cisgender	0.994	0.992	-0.002	0.552
Heterosexual	0.926	0.926	0.000	0.995
Income: More than $1,500	0.627	0.597	-0.031	0.050**
Education				
High School	0.117	0.128	0.011	0.306
More than High School	0.880	0.871	-0.009	0.383
Employed	0.852	0.847	-0.005	0.674
Management Experience	0.771	0.767	-0.004	0.773
Region: Non-Metro	0.366	0.355	-0.011	0.477
Political Affiliation				
Leans Right	0.460	0.454	-0.006	0.704
Leans Left	0.540	0.546	0.006	0.704
Religious Affiliation				
Catholic	0.490	0.483	-0.007	0.644
No Religion	0.376	0.366	-0.011	0.483
LG Comfort	0.975	0.982	0.006	0.172
Has LGB Relative(s) or Friend(s)	0.806	0.798	-0.008	0.552

Notes: Weighted statistics. Treatment A refers to participants randomized into seeing the key sensitive statements in List A as opposed to List B. Treatment B refers to participants randomized into seeing the key sensitive statements in List B as opposed to List A.

Table A2: Instructions Feedback

Indicate the extent to which you agree or disagree with the following statement: "The instructions were clear."	Number of observations	Share of respondents
Totally agree	3,582	89.55
Somewhat agree	314	7.85
Neither agree nor disagree	63	1.57
Somewhat disagree	25	0.62
Totally disagree	16	0.40

Notes: The original Spanish question is as follows: "Indique en qué medida está de acuerdo o en desacuerdo con la siguiente afirmación: 'Las instrucciones fueron claras' ".

Appendix B. Description of the Variables

All respondents were required to provide an answer to all key questions before being allowed to proceed further. As noted below, some supplementary questions allowed the respondents to select options such as "do not know" or "prefer not to answer." The summary statistics for all variables are provided in Table B1.

† Denotes questions with similar or exact language sourced from the 2024 Census of Chile questionnaire (Cuestionario de Viviendas Particulares, Censo de Poblacion y Vivienda 2024).

‡Denotes questions with similar or exact language sourced from the 2018 World Values Survey (WVS7 Questionnaire Chile 2018).

§Denotes questions with similar or exact language sourced from the 2020 Latinobarometro Questionnaire.

- *Age*† reports the respondent's age in years.
- *Indigenous*† indicates whether the respondent belongs or identifies as belonging to an indigenous or native people group.
- *Race*† reports whether the respondent, according to their ancestors, traditions and culture, is or considers themselves to be of African descent/ancestry.
- *Marital Status*† reports the respondent's current marital status.
- *Household Composition*‡ reports the number of people who live as members of the respondent's household, including children.
- *Children*‡ indicates whether the respondent has children.
- *Education*† reports the highest level of education attained by the respondent.
- *Employment*† reports the employment status of the respondent during the previous week.
- *Employment Role*† reports the role (e.g., employer, employee, domestic worker) of the respondent (if employed) during the previous week.
- *Employment Management* indicates whether the respondent has had any work experience as a supervisor of one or more workers
- *Employment Contract* indicates whether the respondent has a written contract in their main job or economic activity.
- *Employment Report* indicates whether the respondent normally reports income from their job or main economic activity to any government entity.
- *Occupation*§ reports the occupational group of the respondent in general, independent of whether the respondent was employed the previous week.
- *Number (of) Colleagues* reports the number of people (including colleagues, bosses, and employees) that the respondent typically interacts with in a week in their current job (if employed).
- *Region* reports the region of Chile in which the respondent currently lives.
- *Commune* reports the commune in which the respondent currently lives (only if the participant indicated the Metropolitan Region of Santiago in the region question).
- *Birth Country* reports the country of birth of the respondent.
- *Sex*† indicates the sex (male/female) of the participant.
- *Gender Identity*† reports the gender with which the respondent identifies (includes the don't know/prefer not to answer options).

- *Sexual Identity* reports the sexual orientation the respondent identifies with (includes the don't know/prefer not to answer options).
- *Religion*[†] reports the respondent's religion (includes the prefer not to answer option).
- *Political Affiliation*[‡] reports the respondent's political affiliation on a 1-10 scale, with 1 denoting "left" and 10 denoting "right" (response is not required).
- *LG Comfort* takes a value of 1 if the respondent answered positively to all the direct questions about being comfortable with a gay or lesbian individual in the survey. Individuals were asked the following: "Would you feel comfortable… [having a gay boss] / [having a gay dentist] / [having to work closely with a lesbian co-worker] / [having a gay neighbor] / [having a gay real estate agent] / [having a gay taxi driver] / [being served by a gay waiter]?"
- *LGB Relatives or Friends* takes a value of 1 if the respondent reported knowing someone who is lesbian, gay, or bisexual among their immediate family, relatives, neighbors, co-workers, or friends.

B.1 Donation lottery

After the survey module of this study, individuals were asked whether they would like to donate some or all of their potential winnings from a CLP 100,000 (approximately USD 100) raffle to a local LGBTQ-related NGO, as described in Section 3.5. The two local NGOs, "Movilh" and "Iguales," were chosen in consultation with the local partner, Datavoz, and they were each randomly shown to roughly half of the sample (1,986 respondents could donate their winnings to Movilh, and 2,014 participants could donate their winnings to Iguales).

The respondents were shown the following descriptions about these organizations: for Movilh, "The Homosexual Integration and Liberation Movement (Movilh, in Spanish) has been an organization defending the human rights of lesbians, gays, bisexuals, and trans and intersex (LGBTI) individuals since June 28, 1991, with interventions on a national scale covering the social, cultural, political, economic, legal, and legislative spheres;"; for Iguales, the organization "Iguales" works "to achieve the full inclusion of sexual diversity in Chilean society. To this end, they participate in all stages of the formulation of public policies at both the legislative and administrative levels." These mission statements are publicly available on the organizations' websites.

After the end of the study, Datavoz contacted both NGOs to attempt to make the combined donations of all participants. Donations to Iguales were made successfully. Unfortunately, Movilh was unable to accept the donations and provide the necessary tax-related paperwork to Datavoz. For this reason, Datavoz contacted the participants who made donations to Movilh, explained the logistical barrier to making the donations, and returned the donated amounts to the participants as additional bonuses.

Table B1. Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Age	4,000	45.93	14.06	18	87
Between 18 and 34	4,000	0.320	0.467	0	1
Between 35 and 49	4,000	0.274	0.446	0	1
Between 50 and 64	4,000	0.295	0.456	0	1
65 or older	4,000	0.111	0.314	0	1
Sex: Female	4,000	0.511	0.500	0	1
African Descent	4,000	0.049	0.215	0	1
Indigenous	4,000	0.101	0.301	0	1
Gender Identity and Sexual Orientation					
Cisgender	3,990	0.991	0.094	0	1
Heterosexual	3,952	0.920	0.271	0	1
Income: More than $1,500	3,838	0.388	0.487	0	1
Education					
High School	4,000	0.316	0.465	0	1
More than High School	4,000	0.677	0.468	0	1
Employed	4,000	0.806	0.395	0	1
Management Experience	4,000	0.667	0.471	0	1
Region: Non-Metro	4,000	0.578	0.494	0	1
Political Affiliation					
Leans Right	3,763	0.445	0.497	0	1
Leans Left	3,763	0.555	0.497	0	1
Religious Affiliation					
Catholic	3,899	0.499	0.500	0	1
No Religion	3,899	0.320	0.467	0	1
LG Comfort	3,742	0.976	0.153	0	1
Has LGB Relative(s) or Friend(s)	4,000	0.777	0.417	0	1

Notes: Weighted statistics. Missing observations reflect cases in which the respondents did not respond to a question or selected options such as “I prefer not to respond.”

Appendix C. Prolific pilot.

To pilot the list experiments and survey, individuals from Chile were recruited on Prolific before Datavoz started any data collection. The Qualtrics questionnaire used with the Prolific participants is the same as that used with the Datavoz participant and is reported in Appendix E.3, with a few important exceptions. First, as analyzed in Section C.1, only half of the participants saw the list experiments, while the other half only took part in the survey. Second, as analyzed in Section C.2, the Prolific participants were asked additional questions to ensure that the language used in the Datavoz questionnaire was up to date and appropriate and to better interpret the answers to some of the key questions. Third, as discussed in Section C.3, the Prolific participants were asked additional questions after the list experiments to test whether they thought that the instructions were clear and whether they had guessed the topic of the study. In addition, to verify that this study actually measured taste-based rather than statistical discrimination, the Prolific participants were asked to explain and motivate their answers to all three direct questions to test whether individuals' expressions of comfort or discomfort with gay men were driven by prejudice or preferences rather than by productivity expectations or beliefs (Section C.4).

Prolific has been used in many studies in economics (Zmigrod, Rentfrow, and Robbins 2018; Schild et al. 2019; Isler, Maule, and Starmer 2018; Oreffice and Quintana-Domeque 2021; Aksoy, Carpenter, and Sansone 2023; 2025). Available evidence indicates some important advantages of Prolific over the Amazon Mechanical Turk for conducting research: Prolific participants are more diverse, less dishonest, pay more attention to study instructions, and produce higher-quality data (Peer et al. 2017; Palan and Schitter 2018; Peer et al. 2021; Gupta, Rigotti, and Wilson 2021).

This pilot collected responses from 535 participants from Chile between 18 July and 31 July 2024. The main study took approximately 11 minutes and 45 seconds, on average, to complete, and participants who successfully completed the study received USD 3 which, on average, corresponds to USD 15.3/hour.

C.1 Effect of list experiments on direct questions

One may worry that the participants' responses to the key direct questions may have been affected by their answers to the list experiments. For instance, it is possible that the participants wanted to be consistent in their answers and report not being comfortable supervising a gay employee if this was in line with their responses to the related list experiment, even if they may have reported feeling comfortable supervising a gay employee if they had not seen the list experiment beforehand and they had felt social pressure to answer affirmatively.

This channel is somewhat unlikely since the respondents saw a total of 27 items in the list experiments and were asked the related direct questions for only three of these items. Moreover, instead of asking the direct questions right after their corresponding lists, as done, for example, in (Coffman, Coffman, and Ericson 2017), the direct questions were located after the demographic questions and together with other questions on income, religiosity, and political affiliation, in line with (Lax, Phillips, and Stollwerk 2016), (Chuang et al. 2021), and (Aksoy,

Carpenter, and Sansone 2025). This choice should reduce the probability of respondents linking the direct questions to the related list experiment.

In addition, while (Lax, Phillips, and Stollwerk 2016) asked the direct questions after their list experiment on support for same-sex marriage, they used a single experiment method, which made it possible for them to study the impact of seeing the key item in the list experiment on the direct survey question responses. Reassuringly, they did not find any significant impact coming from the fact that half of their subjects saw the key statement twice (once in the list experiment and once as a direct question).

To produce a similar test, only half of the respondents in the Prolific pilot were asked to respond to both the list experiments and the survey questions, while the other half were shown only the survey questions (including the direct questions). As shown in Table C1, the average levels of comfort with gay individuals, as reported in the direct questions, were not affected by whether the respondents were also asked to participate in the list experiments before answering the direct questions.

C.2 Terminology

The pilot on Prolific was also leveraged to ensure that the language used in the main questionnaire was up to date and appropriate. Indeed, the respondents were asked which term they thought was more commonly used to refer to a person working with them. As shown in Table C2, "co-worker" was the most frequently selected term. Therefore, this terminology was adopted in the main Datavoz survey.

In addition, the respondents were asked which term they thought was more commonly used to refer to people who are attracted to individuals of the same sex. As shown in Table C3, "gay" was the most frequently selected term. Therefore, this terminology was adopted in the main Datavoz survey.

The respondents were then asked whether they thought that one is talking about gay men or lesbian women when one refers to gay or homosexual individuals. As shown in Table C4, most respondents thought that "gay"/"homosexual" referred mainly to both gay men and lesbian women. This finding is important for interpreting, from a gender perspective, the main findings from the list experiments. Nevertheless, since Spanish is a gendered language, the main three key statements in the list experiments implicitly treat gay men – rather than lesbian women – as the reference category.

C.3 Feedback on survey instructions and the survey aim

As in the main survey, the participants in the Prolific pilot were also administered questions to provide feedback regarding the instructions and the instrument overall. Table C5 tabulates the responses to the following prompt: "I believe that the instructions were clear." All participants, except 4, agreed with the prompt. This question was presented to all participants at the end of the survey.

The subset of participants in the Prolific pilot who were randomized into the list experiment were also provided with questions to leave feedback for the researchers immediately after

completing the list experiment and before continuing with the survey. Specifically, an open-ended question was deployed to ask the respondents what they believed the study was about. Among the top 30 most frequent words in response to this question were "social issues," "opinions," "people," "Chileans," "population," "politics," "beliefs," and "ideas" (translated from Spanish). LGBTQ-related terms[17] were mentioned by only 9 out of 266 participants. Additionally, a manual inspection of all the non-empty open-ended responses confirms that most individuals were not sure or were confused about the purpose of the study, expressed curiosity about the goals of the researchers, and provided guesses along the lines of "eliciting beliefs and opinions from Chileans about controversial social and political issues." These findings confirm that the participants did not believe that the list experiments were focused on measuring attitudes toward sexual minorities, thus achieving the goal of not priming or influencing the respondents.

C.4 Follow-up questions after the direct questions

This study focuses on measuring sources of taste-based discrimination. A key goal of the Prolific pilot was to test whether taste-based preferences (as opposed to statistical discrimination) did in fact explain the respondents' behaviors. To that end, the direct questions (i.e., "Would you feel comfortable supervising a gay employee?," "Would you feel comfortable working closely with a gay co-worker?," and "Would you feel comfortable having a cashier at the supermarket who is gay?") were immediately followed by mandatory open-ended text questions asking the participants to explain their previous yes/no responses.

To analyze these data, an indicator variable was created to flags respondents who justified their responses to any of the three statements with mentions of "productivity," "performance," "output," "efficacy," "efficiency," "stereotype," and "results" (translated from Spanish). All the flagged responses expressed indifference to these qualifiers (i.e., that gay people do not differ from others in regard to these considerations). Restricting the sample of open-ended explanations to the small subsets of respondents who answered negatively to any of the direct questions reveals that most justifications were taste based. For example, some participants expressed discomfort with the idea of a gay employee or coworker insinuating attraction to them. Other participants mentioned that gay people contradict the conservative values that their families raised them with. Only one participant mentioned a stereotype consisting of gay people being "conflictive" and "gossiping": the respondent argued that this presents a risk at work due to the "current laws and political system." Overall, these findings confirm that the list experiments measured potential sources of taste-based discrimination, not statistical discrimination.

[17] LGBTQ-related terms that were searched for include the following: "lgbt," "lgbtq," "lgbtq+," "LGBT," "LGBTQ," "LGBTQ+," "gay," "lesbiana," "transgénero," "bisexual," "queer," "homosexual," "transexual," "género," "orientación sexual," and "diversidad sexual."

Table C1. Effect of List Experiments on Direct Questions

	With list experiments	Direct questions only	Difference
Would you feel comfortable…	(1)	(2)	(2) – (1)
supervising a gay employee?	0.96	0.97	0.01
	(0.20)	(0.18)	[0.6191]
working closely with a gay colleague?	0.95	0.96	0.01
	(0.21)	(0.20)	[0.8032]
having a cashier at the supermarket who is gay?	0.99	0.98	-0.01
	(0.09)	(0.15)	[0.1610]

Notes: Column 1 reports the average responses to the three direct questions among participants who also responded to the list experiments. Column 2 reports the average responses to the three direct questions among participants who were not shown the list experiments. Column 3 reports the difference between the two groups. The standard deviation is reported in parentheses in Columns 1 and 2. P-values are reported in square brackets in Column 3.

Table C2. Colleague versus Co-worker

Which term do you think is most commonly used to refer to a person who works with you?	Number of observations	Share of respondents
Colega	181	33.83
Compañero de trabajo	329	61.50
Don't know/not sure	25	4.67

Notes: The original Spanish question is as follows: "¿Qué término crees que es más común para referirse a una persona que trabaja contigo?".

Table C3. Gay versus Homosexual

Which term do you think is most commonly used to refer to a person who is attracted to people of the same sex?	Number of observations	Share of respondents
Gay	356	66.54
Homosexual	142	26.54
Don't know/not sure	37	6.92

Notes: The original Spanish question is as follows: "¿Qué término crees que es más común para referirse a una persona que siente atracción hacia personas del mismo sexo?".

Table C4. Gay/Homosexual and Gender

When we refer to gay/homosexual people, do you think we mean...?	Number of observations	Share of respondents
Both	398	74.39
Gay/homosexual men	132	24.67
Gay/homosexual/lesbian women	1	0.19
Don't know/not sure	4	0.75

Notes: The original Spanish question is as follows: "Cuando nos referimos a personas gays/homosexuales, ¿crees que nos referimos a...?".

Table C5. Instructions Feedback – All Prolific Pilot Participants

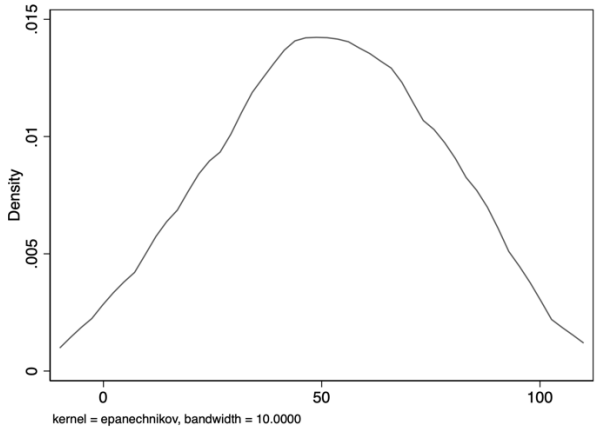
Indicate the extent to which you agree or disagree with the following statement: "The instructions were clear."	Number of observations	Share of respondents
Totally agree	517	96.64
Somewhat agree	14	2.62
Neither agree nor disagree	0	0.19
Somewhat disagree	1	0
Totally disagree	3	0.56

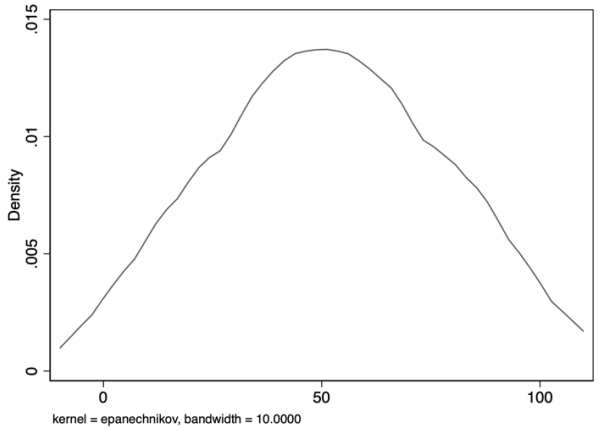
Notes: The original Spanish question is as follows: "Indique en qué medida está de acuerdo o en desacuerdo con la siguiente afirmación: "Las instrucciones fueron claras"."

Appendix D. Additional Tables and Figures

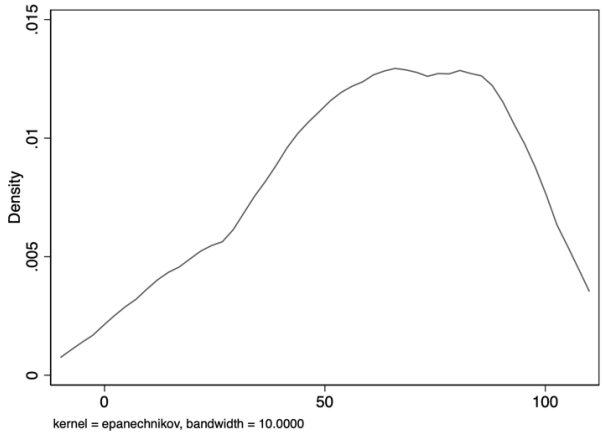
Figure D1. Perceptions of General Views on Attitudes toward Gay People – Kernel Densities

Panel A: Supervisors **Panel B: Co-workers**



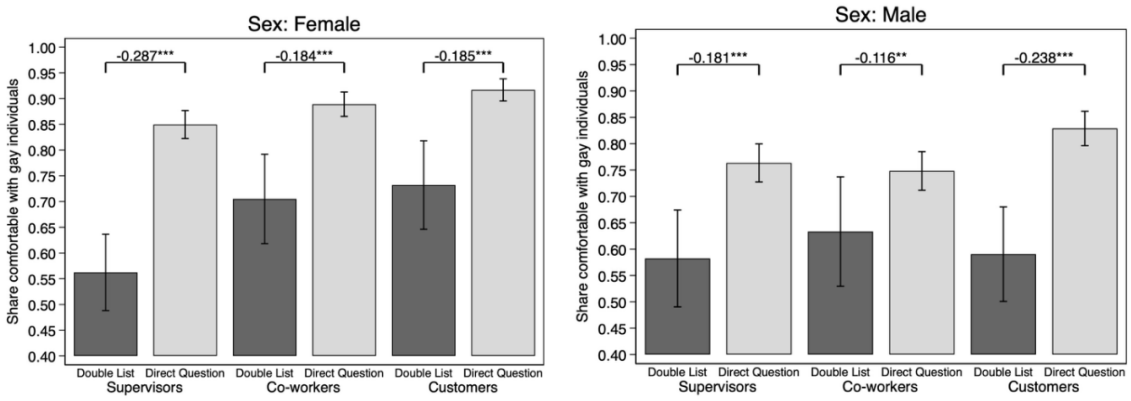


Panel C: Customers



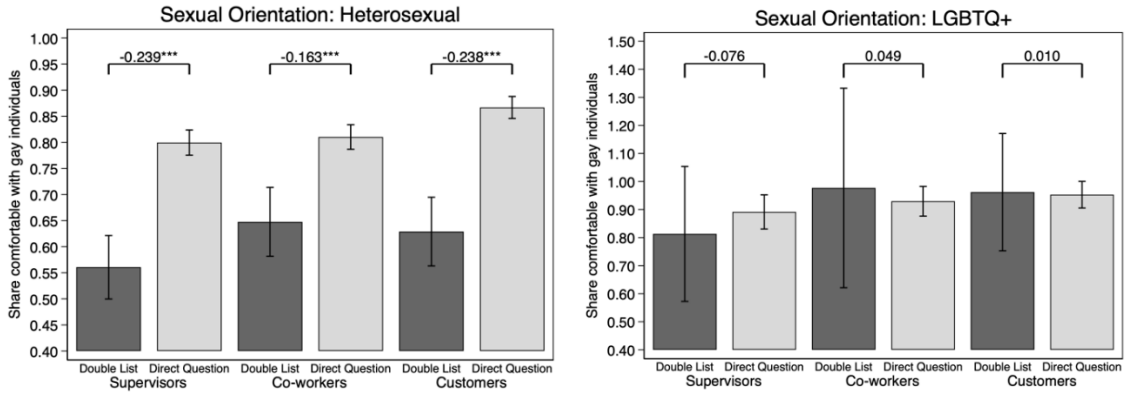
Notes: Weighted statistics. Number of observations: 4,000. See also the notes in Figure 2.

Figure D2. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Sex



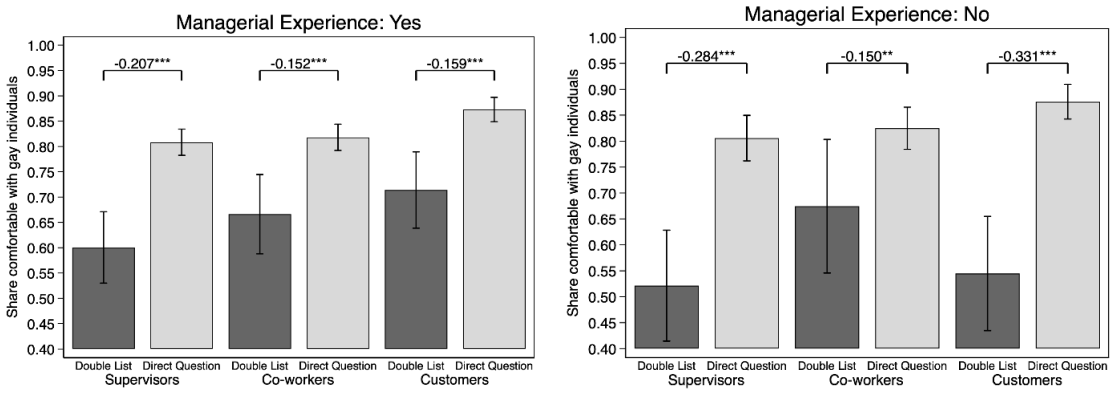
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 1,879 (sex: female) and 2,121 (sex: male). *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D3. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Sexual Orientation



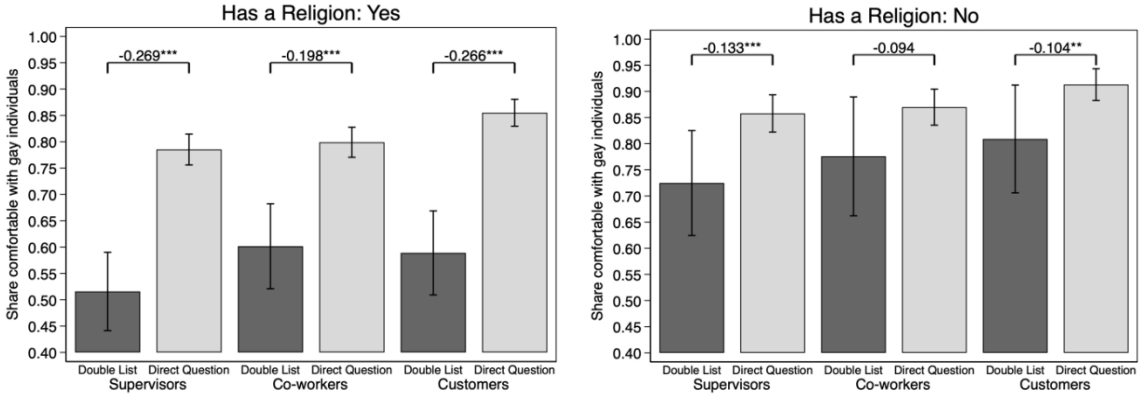
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,659 (heterosexual) and 293 (LGBTQ+). The sexual orientation question is translated from the original Spanish as follows: "Currently, you identify as… *{presented in random order}* [gay (attraction of a man to another man) / lesbian (attraction of a women to another woman / bisexual (attraction to more than one sex or gender) / heterosexual (attraction to the opposite sex) / I use a different term (please specify) / I don't know / I prefer not to answer]." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D4. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Managerial Experience



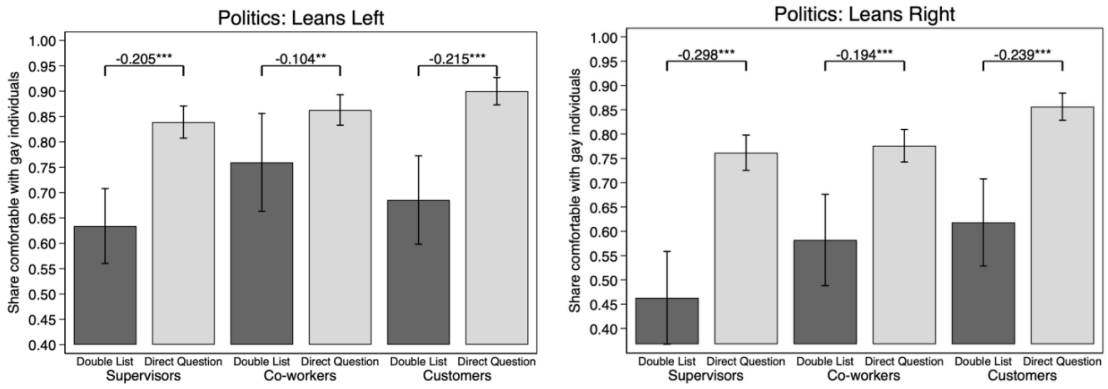
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 3,076 (managerial experience: yes) and 924 (managerial experience: no). The managerial experience question is translated from the original Spanish as follows: "Have you ever had any work experience as a supervisor of one or more workers? [Yes / No]" *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D5. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Religiosity



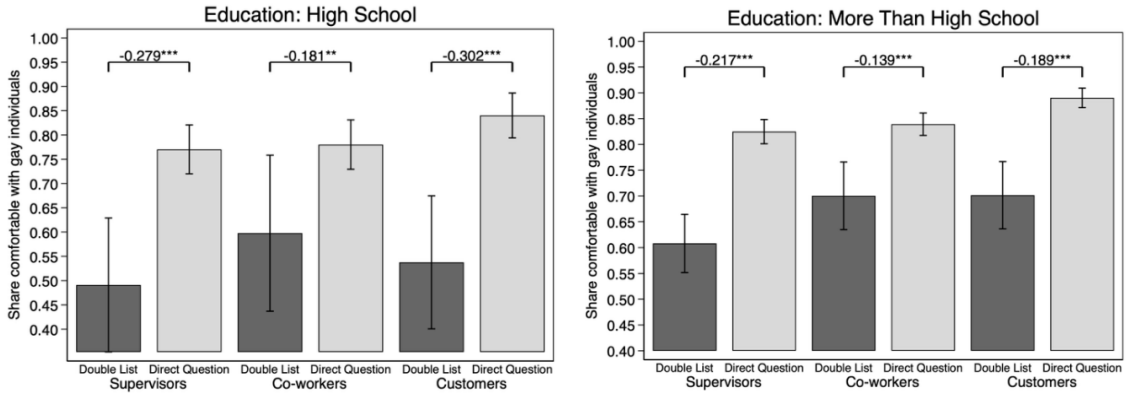
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 2,456 (has a religion: yes) and 1,443 (has a religion: no). The religion question is translated from the original Spanish as follows: "What is your religion or creed? [Catholic / Evangelical or Protestant / Jewish / Muslim / Mormon / Orthodox Catholic / Buddhist / Hindu / Bahá'í Faith / Jehovah's Witness / Other (please specify) / None / I prefer not to answer]." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D6. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Political Affiliation



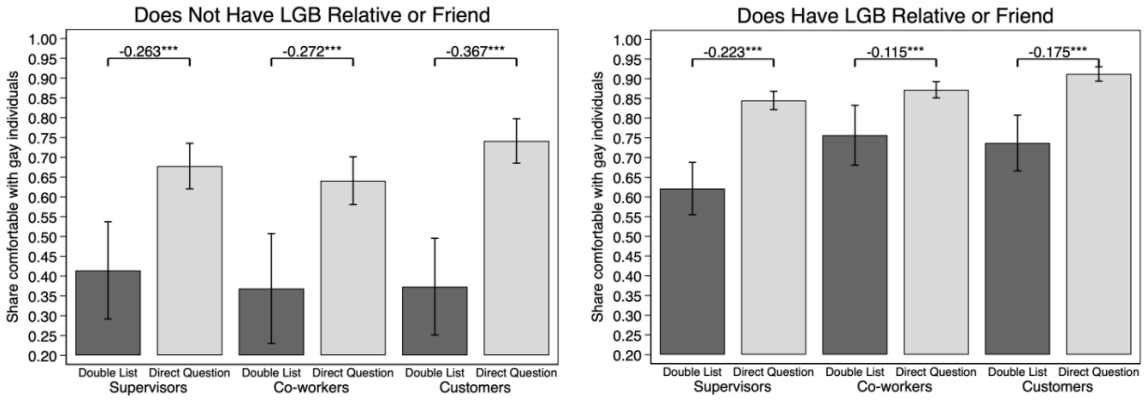
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 2,043 (politics: leans left) and 1,720 (politics: leans right). The politics question is translated from the original Spanish as follows: "In political matters, people talk about "the left" and "the right." Where on this scale, where 1 is left and 10 is right, would you place yourself? *{scale where 1 was placed as the leftmost option and 10 was placed as the rightmost option was presented}* / I don't know / I prefer not to answer]." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D7. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Education



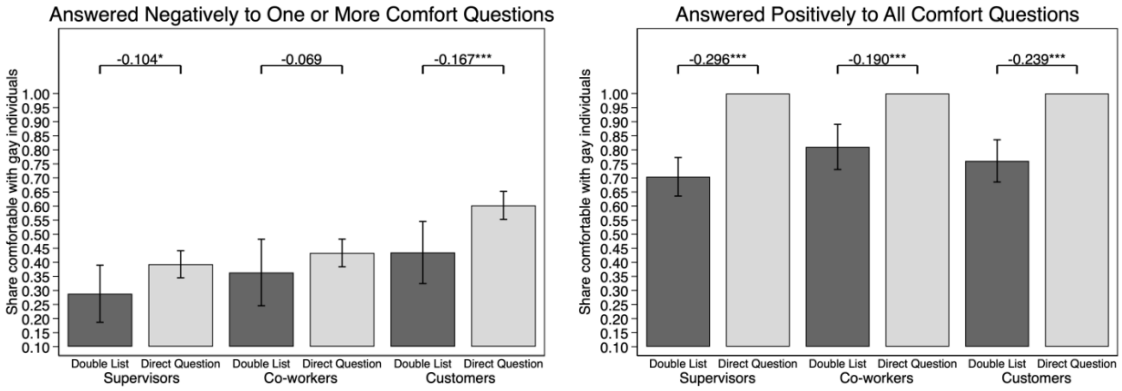
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 490 (education: high school) and 3,501 (education: more than high school). The education question is translated from the original Spanish as follows: "What is your highest level of education achieved? (A person achieved or reached an educational level when they declare having completed at least one course of the corresponding level) [Never attended / Nursery / Childcare center / Pre-K / Kindergarten / Special or preferential education / Basic education / Primary (old system) / Scientific, humanities, or artistic secondary education / Technical secondary education / Humanities (old system) / Commercial, industrial, or pedagogical (old system) / Higher level technician (1 to 3 years) (including sub-officer in the Armed Forces) / Professional (4 years or more) (including officer in the Armed Forces) / Master's degree / PhD." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D8. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by LGB Familiarity



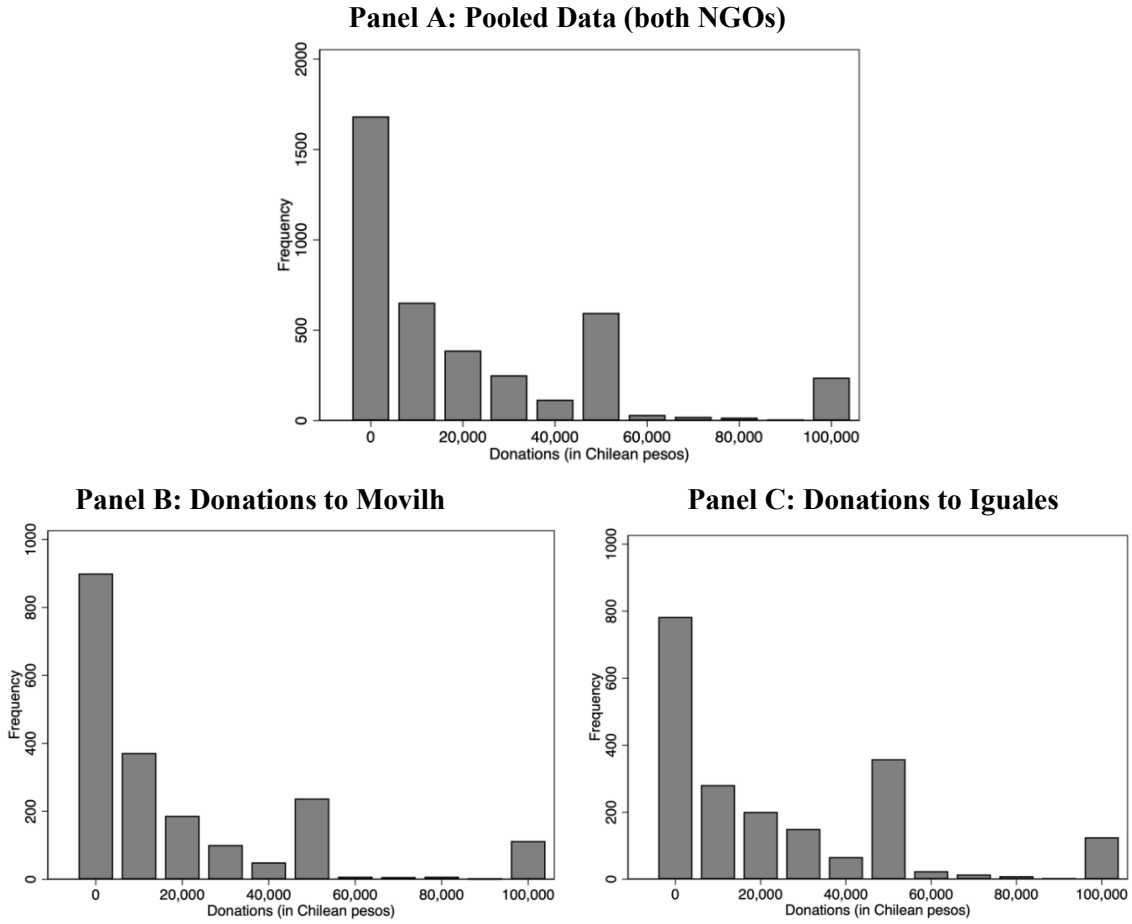
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 793 (does not have an LGB relative or friend) and 3,207 (does have an LGB relative or friend). The question is translated from the original Spanish as follows: "Among your immediate family members, relatives, neighbors, coworkers, or close friends, are any of them gay, lesbian, or bisexual (that you know of)? [Yes / No]." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D9. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Additional Comfort Questions on Lesbian/Gay Individuals



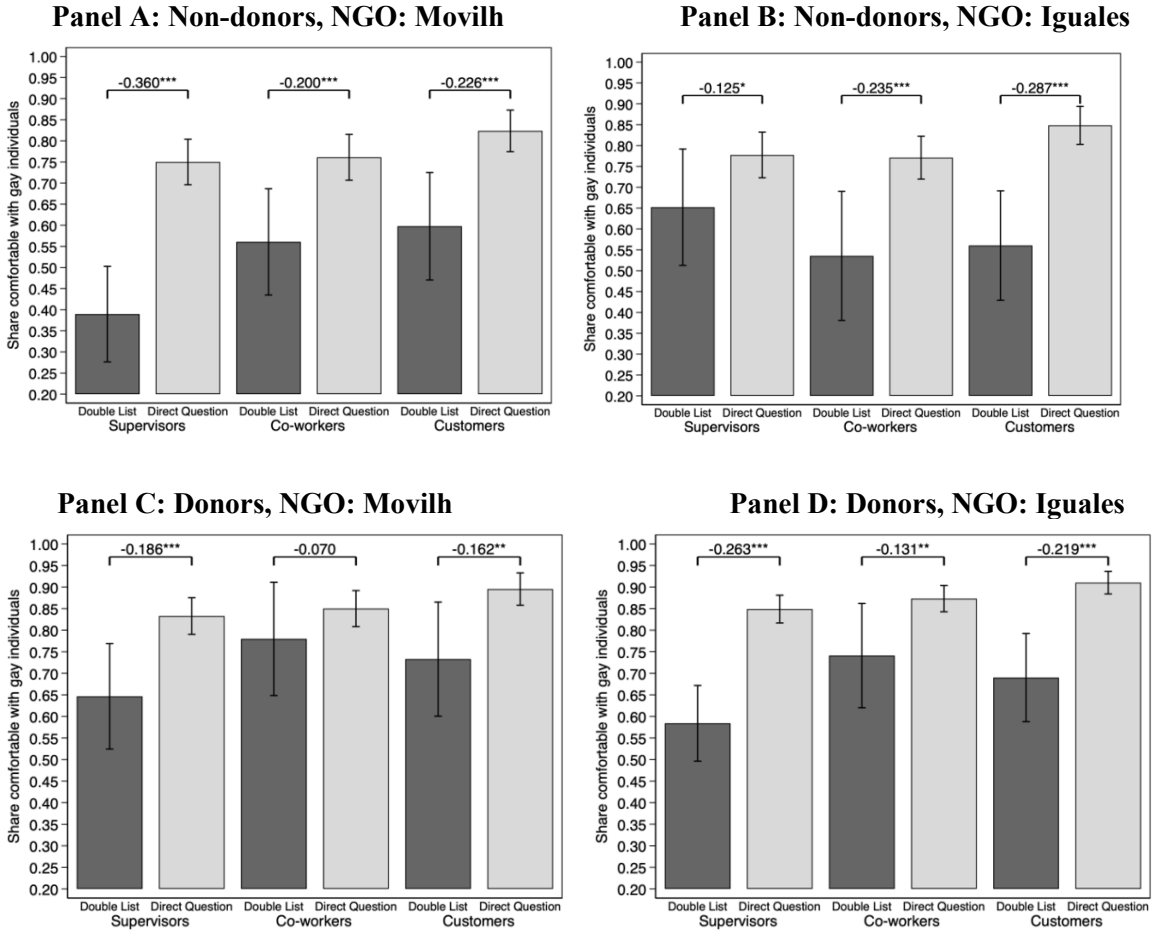
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 1,173 (answered negatively to one or more comfort questions) and 2,827 (answered positively to all comfort questions). The question is translated from the original Spanish as follows: "Would you feel comfortable… [having a gay person as a neighbor / having a gay dentist / being served by a gay waiter / dealing with a gay real estate agent / having a gay boss / having to work closely with a lesbian colleague / having to work closely with an indigenous colleague / having a gay taxi driver / supervising several workers]? [Yes / No]." *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D10. Histogram of Donation Decisions



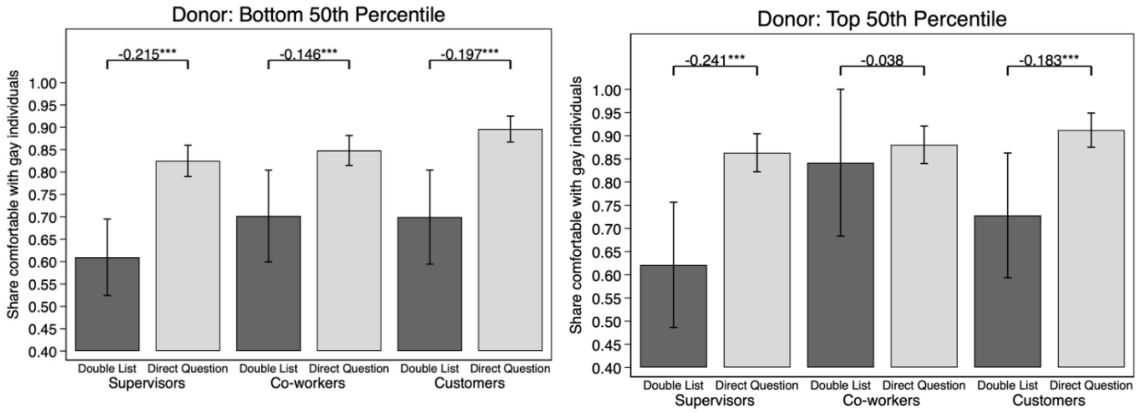
Notes: Unweighted statistics. Number of observations: 4,000 (Panel A), 1,986 (Panel B), and 2,014 (Panel C).

Figure D11. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Donor Status and NGO



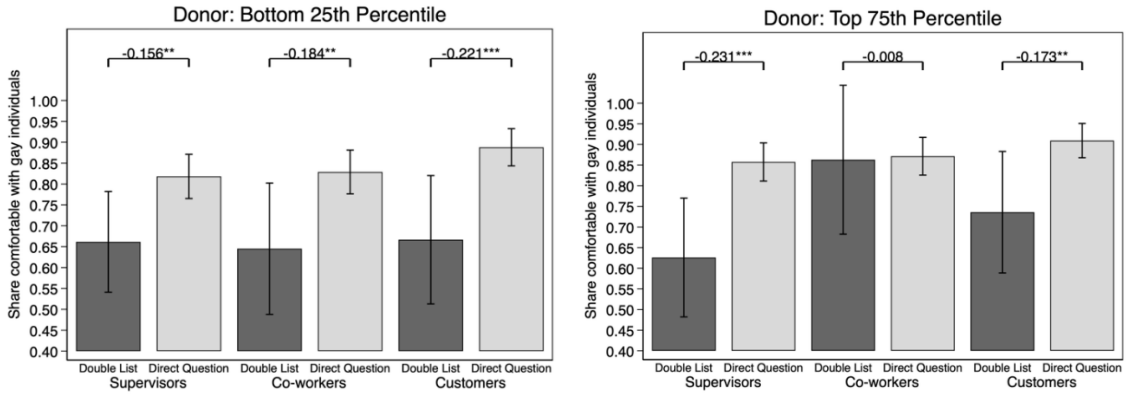
Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Panel A: The sample is restricted to individuals who chose to not donate any amount of the raffled monetary prize, and their randomly assigned NGO was Movilh (N= 900). Panel B: The sample is restricted to individuals who chose to not donate any amount of the raffled monetary prize, and their randomly assigned NGO was Iguales (N= 783). Panel C: The sample is restricted to individuals who chose to donate a positive amount of the raffled monetary prize, and their randomly assigned NGO was Movilh (N= 1,086). Panel D: The sample is restricted to individuals who chose to donate a positive amount of the raffled monetary prize, and their randomly assigned NGO was Iguales (N= 1,231). Total number of observations: 4,000. *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D12. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Top and Bottom 50th Percentile of Donors



Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 1,292 (donor: bottom 50th percentile; donation <= CLP 30,000) and 1,025 (donor: top 50th percentile; donation > CLP 30,000). *p < 0.10; **p < 0.05; ***p < 0.01.

Figure D13. List Experiment on Supervisors', Co-workers', and Customers' Attitudes – Heterogeneity by Top and Bottom 75th/25th Percentile of Donors



Notes: Weighted statistics. 95-percent confidence intervals are reported with the vertical range plots. The numbers above the horizontal bars are the differences between the two groups at the base of each horizontal bar. See also the notes in Figure 1. Number of observations: 653 (donor: bottom 25th percentile; donation <= CLP 10,000) and 909 (donor: top 75th percentile; donation >= CLP 50,000). *p < 0.10; **p < 0.05; ***p < 0.01.

Table D1. List Experiments – Differences in Means Comparisons and Robustness Checks

	List A	List B	Double List	Direct Question	(1)-(2)	(4)-(3)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Main Data (Weighted)						
Supervisor	0.676	0.472	0.574	0.807	0.204**	0.234***
	(0.054)	(0.050)	(0.030)	(0.012)	(0.085)	(0.030)
Co-worker	0.772	0.567	0.669	0.820	0.205**	0.151***
	(0.054)	(0.046)	(0.034)	(0.011)	(0.073)	(0.034)
Customer	0.685	0.628	0.657	0.874	0.057	0.217***
	(0.054)	(0.055)	(0.032)	(0.010)	(0.088)	(0.032)
Panel B: Main Data (Unweighted)						
Supervisor	0.649	0.604	0.627	0.827	0.045	0.200***
	(0.024)	(0.025)	(0.016)	(0.006)	(0.036)	(0.016)
Co-worker	0.693	0.644	0.668	0.838	0.049	0.170***
	(0.025)	(0.024)	(0.016)	(0.006)	(0.036)	(0.016)
Customer	0.696	0.761	0.728	0.893	-0.065	0.165***
	(0.025)	(0.027)	(0.017)	(0.005)	(0.038)	(0.017)
Panel C: Main Data and Pilots (Unweighted)						
Supervisor	0.621	0.581	0.621	0.828	0.040	0.206***
	(0.016)	(0.016)	(0.016)	(0.006)	(0.034)	(0.016)
Co-worker	0.669	0.623	0.669	0.837	0.046	0.168***
	(0.016)	(0.016)	(0.016)	(0.006)	(0.034)	(0.016)
Customer	0.723	0.670	0.723	0.893	-0.053	0.170***
	(0.017)	(0.037)	(0.032)	(0.005)	(0.037)	(0.017)

Notes: Standard errors are in parentheses. Columns 1 and 2 report the mean differences in responses across treatments in List A and List B, respectively. Column 3 reports the share of individuals comfortable with the corresponding key statement elicited by the double list experiment. Column 4 reports the mean of the direct question for each corresponding statement. Column 5 reports the differences between Columns 1 and 2, and Column 6 reports the differences between Columns 4 and 3. Supervisor key statement: "I would feel comfortable supervising a gay employee." Co-worker key statement: "I would feel comfortable working closely with a gay co-worker." Customer key statement: "I would feel comfortable having a cashier at the supermarket who is gay." Number of observations: 4,000 (Panels A and B) and 4,297 (Panel C). *p < 0.10; **p < 0.05; ***p < 0.01.

Table D2. List Experiment on Supervisors', Co-workers', and Customers' Attitudes toward Gay Individuals – Heterogeneity Analyses with Additional Controls

Interaction of treatment variable with:	Supervisors (1)	Co-workers (2)	Customers (3)
Age: 18–44	0.103	0.097	-0.072
	(0.067)	(0.071)	(0.066)
Race: African Descent	-0.333*	0.001	0.084
	(0.202)	(0.176)	(0.204)
Indigenous	0.001	0.128	-0.062
	(0.112)	(0.124)	(0.126)
Sex: Female	-0.031	0.102	0.148**
	(0.062)	(0.071)	(0.067)
Sexual orientation: Heterosexual	-0.162	-0.205	-0.330***
	(0.115)	(0.177)	(0.117)
Household income: More than $1,500	0.064	-0.193***	-0.019
	(0.060)	(0.065)	(0.068)
Education: More than high school	0.013	0.051	0.133*
	(0.080)	(0.088)	(0.080)
Employment status: Employed	0.086	0.118	-0.057
	(0.083)	(0.079)	(0.086)
Management Experience	0.129*	0.104	0.140**
	(0.071)	(0.079)	(0.068)
Region: Outside metro	0.021	-0.089	-0.044
	(0.058)	(0.064)	(0.061)
Political affiliation: Lean left	0.053	-0.024	-0.033
	(0.069)	(0.076)	(0.068)
Current religious affiliation: Not religious	0.095	0.086	0.199***
	(0.067)	(0.078)	(0.073)
LGB Comfort	0.380***	0.332***	0.245***
	(0.076)	(0.080)	(0.077)
Has LGB Relative(s) and/or Friend(s)	0.029	0.229**	0.179**
	(0.088)	(0.091)	(0.082)
Belief: 50% or more comfortable supervising gay employees	0.066		
	(0.061)		
Belief: 50% or more comfortable with gay co-workers		0.131*	
		(0.068)	
Belief: 50% or more comfortable with gay cashiers			0.018
			(0.070)
Constant	0.125	0.198	0.442**
	(0.172)	(0.223)	(0.180)
Observations	3,518	3,518	3,518

Notes: Heterogeneity weighted multivariate analysis. See also the notes in Table 3. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Appendix E. Datavoz questionnaire

E.1 Recruitment emails

E.1.1 Original Spanish Version of Recruitment Email

ASUNTO: Invitación a participar en encuesta BID-DATAVOZ

Datavoz por encargo del Banco Interamericano del Desarrollo (BID) se encuentra realizando una encuesta. La encuesta implica responder a preguntas bajo diferentes escenarios según sus preferencias individuales.

La participación debería tomar aproximadamente 10 minutos.

Se sortearán 50 tarjetas de regalo (gift cards) de $50.000 pesos chilenos entre los participantes que completen la encuesta. Además, un subconjunto aleatorio de 50 encuestados podrá recibir $100.000 pesos chilenos adicionales en forma de gift cards.

<<ENLACE DE ENCUESTA>>

Muchas gracias

E.1.2 English Version of Recruitment Email

SUBJECT LINE: Invitation to participate in survey IDB-DATAVOZ

Datavoz, commissioned by the Inter-American Development Bank (IDB), is conducting a survey. The survey involves answering questions under different scenarios based on your individual preferences.

Participation should take approximately 10 minutes.

Fifty gift cards worth 50,000 Chilean pesos each will be raffled among the participants who complete the survey. Additionally, a random subset of 50 respondents may receive an extra 100,000 Chilean pesos in the form of gift cards.

<<**LINK TO SURVEY**>>

Thank you very much.

E.2 Consent Forms

E.2.1 Original Spanish Version of Consent Form

Institutional Review Board
1204 Marie Mount Hall ● 7814 Regents Drive ● College Park, MD 20742 ● 301-405-4212 ● irb@umd.edu

CONSENTIMIENTO PARA PARTICIPAR

Título del Proyecto	Encuesta de Opiniones sobre Cuestiones Sociales en Chile
Investigador Principal	Esta investigación es realizada por Ariel Listo de la Universidad de Maryland, College Park, Ercio Munoz del Banco Interamericano de Desarrollo (BID), y Dario Sansone de la Universidad de Exeter, en colaboración con DATAVOZ.
Propósito del Estudio	El propósito de este proyecto de investigación es entender las opiniones de las personas sobre ciertas cuestiones sociales.
Procedimientos	La encuesta implica responder a preguntas bajo diferentes escenarios según sus preferencias individuales. En algunos casos, sus decisiones pueden tener un efecto real en sus ganancias. La participación debería tomar aproximadamente 10 minutos.
Potenciales Riesgos y Molestias	No existen riesgos físicos, psicológicos o sociales conocidos para los participantes aparte de la ansiedad o el aburrimiento asociados con la participación en este proyecto de investigación. Puede tomar descansos según sea necesario.
Beneficios Potenciales	Esta investigación no está diseñada para ayudarlo personalmente, pero los resultados pueden ayudar a los investigadores a aprender más sobre las opiniones sociales. Esperamos que en el futuro otras personas puedan beneficiarse de este estudio a través de una mejor comprensión de las preferencias individuales.
Confidencialidad	No accederemos a ninguna información personal identificable sobre usted. Los investigadores no recibirán datos identificadores sobre usted. Todos los datos no identificables se almacenarán en computadoras protegidas por contraseña y se compartirán solo a través de servicios en línea seguros. Si escribimos un informe o artículo sobre este proyecto de investigación, su identidad será protegida al máximo posible. Su información puede ser compartida con representantes de la Universidad de Maryland, College Park o autoridades gubernamentales si usted o alguien más está en peligro o si estamos obligados a hacerlo por ley.
Compensación	Se sortearán 50 tarjetas de regalo (gift cards) de $50.000 pesos chilenos entre los participantes que completen la encuesta. Además, un subconjunto aleatorio de 50 encuestados podrán recibir $100.000 pesos chilenos adicionales en forma de gift cards. Si decide no terminar la encuesta, no entrará en la rifa ni recibirá ningún tipo de compensación parcial. Usted será responsable de cualquier impuesto evaluado sobre la compensación.
Derecho a Retirarse y Preguntas	Su participación en esta investigación es completamente voluntaria. Puede optar por no participar en absoluto. Si decide participar en esta investigación, puede dejar de participar en cualquier momento. Si decide no participar en este estudio o si deja de participar en cualquier momento, no será penalizado ni perderá

	ningún beneficio para el cual de otra manera calificaría, pero no entrará en la rifa ni recibirá ningún tipo de compensación parcial. Si decide dejar de participar en el estudio, si tiene preguntas, inquietudes o quejas, o si necesita informar una lesión relacionada con la investigación, por favor contacte al investigador: Ariel Listo 2200 Symons Hall, University of Maryland, College Park, MD 20742 USA alisto@umd.edu (301) 405-1293 Dr. Dario Sansone University of Exeter Business School, Rennes Drive, Exeter EX4 4PU, Reino Unido Email: d.sansone@exeter.ac.uk
Derechos del Participante	Si tiene preguntas sobre sus derechos como participante en la investigación o desea informar una lesión relacionada con la investigación, por favor contacte: University of Maryland College Park Institutional Review Board Office 1204 Marie Mount Hall College Park, Maryland, 20742 USA E-mail: irb@umd.edu Telephone: 301-405-0678 Para más información sobre los derechos de los participantes, por favor visite: https://research.umd.edu/research-resources/research-compliance/institutional-review-board-irb/research-participants Esta investigación ha sido revisada de acuerdo con los procedimientos de IRB de la Universidad de Maryland, College Park para investigaciones que involucran a sujetos humanos.
Declaración de Consentimiento	Su consentimiento indica que usted tiene al menos 18 años de edad; ha leído este formulario de consentimiento o se lo han leído; y acepta voluntariamente participar en este estudio de investigación. Puede descargar esta hoja informativa para sus registros.

E.2.1 English Translation of Consent Form

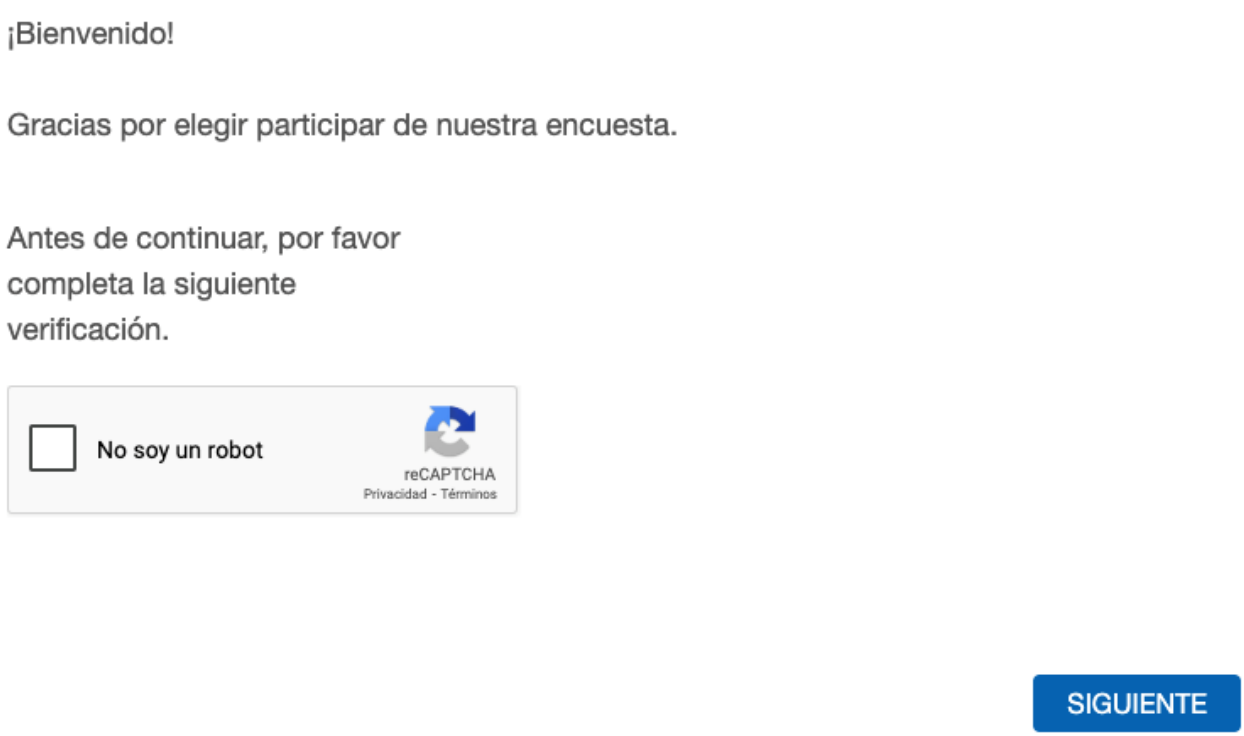
Institutional Review Board
1204 Marie Mount Hall ● 7814 Regents Drive ● College Park, MD 20742 ● 301-405-4212 ● irb@umd.edu

CONSENT TO PARTICIPATE

Project Title	Survey on Opinions about Social Issues in Chile
Principal Investigator	This research is being conducted by Ariel Listo at the University of Maryland, College Park, Ercio Munoz at the Inter-American Development Bank (IDB), and Dario Sansone at the University of Exeter, in collaboration with DATAVOZ.
Purpose of the Study	The purpose of this research project is to understand individuals' views on certain social issues.
Procedures	The survey involves responding to questions and making choices under different scenarios according to your individual preferences. In some cases, your decisions may have a real effect on your earnings. Participation should take about 10 minutes.
Potential Risks and Discomforts	There are no known physical, psychological, or social risks to subjects, other than anxiety or boredom associated with participating in this research project. You may take breaks as needed.
Potential Benefits	This research is not designed to help you personally, but the results may help the investigators learn more about social views. We hope that, in the future, other people might benefit from this study through improved understanding of individuals' preferences.
Confidentiality	We will not be accessing any personally identifying information about you. No identifiers will be shared with the investigators. All non-identifiable data will be stored on password-protected computers and shared only through secure online services. If we write a report or article about this research project, your identity will be protected to the maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities if you or someone else is in danger or if we are required to do so by law.
Compensation	50 gift cards of 50,000 Chilean pesos each will be raffled among the participants who complete the survey. Additionally, a random subset of 50 respondents may receive an additional 100,000 Chilean pesos. Should you choose to not finish the survey, you will not be entered into the raffle or receive any type of partial compensation. You will be responsible for any taxes assessed on the compensation.
Right to Withdraw and Questions	Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify, but you will not be entered into the raffle or receive any type of partial compensation.

	If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the investigators: Ariel Listo 2200 Symons Hall, University of Maryland, College Park, MD 20742 USA Email: alisto@umd.edu (301) 405-1293 Dr. Dario Sansone University of Exeter Business School, Rennes Drive, Exeter EX4 4PU, UK Email: d.sansone@exeter.ac.uk
Participant Rights	If you have questions about your rights as a research participant or wish to report a research-related injury, please contact: University of Maryland College Park Institutional Review Board Office 1204 Marie Mount Hall College Park, Maryland, 20742 USA E-mail: irb@umd.edu Telephone: 301-405-0678 For more information regarding participant rights, please visit: https://research.umd.edu/research-resources/research-compliance/institutional-review-board-irb/research-participants This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.
Statement of Consent	Your consent indicates that you are at least 18 years of age; you have read this consent form or have had it read to you; and you voluntarily agree to participate in this research study. You may download this information sheet for your records.

E.3 Questionnaire



Por favor, revisa la información contenida en el formulario a continuación.

Consentimiento Informado

Haciendo clic en [Sí, consiento...] a continuación, confirmas que tienes al menos 18 años de edad y que consientes participar en el estudio de investigación descrito anteriormente.

Si haces clic en [No, no consiento...], no serás elegible para participar en este estudio y serás redirigido al final de la encuesta.

- Si, consiento a participar en este estudio.
- No, no consiento a participar en este estudio.

SIGUIENTE

Encuesta de Opiniones sobre Cuestiones Sociales en Chile

Introducción

En este estudio, se te pedirá que respondas algunas preguntas. Tus respuestas serán anónimas. El estudio tomará aproximadamente **10 minutos.**

¿Cuánto recibiré de pago por mi participación?

Sorteo por participación: Se sortearán 50 gift cards de $50.000 pesos entre los participantes que completen la encuesta. Este sorteo se realizará luego de que todos los participantes de esta encuesta, finalicen su participación.

Sorteo por bono adicional: Además del sorteo por participación, se realizará un sorteo antes que tu termines la encuesta en la cual un subconjunto aleatorio de 50 encuestados podrán recibir $100.000 pesos adicionales en forma de gift cards.

DATAVOZ utilizará el correo electrónico a través del cual te contactamos para hacerte llegar tu premio si resultaras ganador de uno de estos sorteos.

Preguntas de Verificación de Atención

En cuestionarios como este, a veces hay participantes que no leen cuidadosamente las preguntas y solo hacen clic rápidamente a través de la encuesta. Esto puede comprometer los resultados de estudios de investigación como este. Como prevención, habrá varias preguntas de **Verificación de Atención** a lo largo de esta encuesta. Si no logras completar correctamente dos o más de estas preguntas, es posible que no recibas compensación.

Por favor, haz clic en SIGUIENTE cuando estés listo.

SIGUIENTE

Instrucciones

En las siguientes páginas, se presentarán listas de afirmaciones que pueden ser o no verdaderas para ti. Las afirmaciones serán sobre ti mismo y tus opiniones sobre temas sociales. Nos gustaría saber cuántas de las afirmaciones en cada lista son verdaderas para ti. En estas listas, no estamos preguntando **cuáles** afirmaciones específicas son verdaderas para ti, solo estamos preguntando **cuántas** de ellas son verdaderas para ti. No existen respuestas buenas o malas. Solo nos interesa conocer su opinión y experiencia.

En la página siguiente, te daremos un ejemplo.

Por favor, haz clic en SIGUIENTE cuando estés listo.

SIGUIENTE

Ejemplo

Queremos asegurarnos de que entiendas cómo funciona esto. Aquí tienes un ejemplo:

***

Hay tres (3) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Tengo una camiseta naranja.
- En mi hogar hay al menos cuatro mascotas.
- Reciclo regularmente

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti: ____

***

Supongamos que sí tienes una camiseta naranja y que en tu hogar hay al menos cuatro mascotas. Pero no reciclas. En ese caso, dos de las afirmaciones anteriores son verdaderas para ti. Por lo tanto, indicarías esto ingresando 2 en el cuadro de respuesta.

Ten en cuenta: No estamos preguntando **cuáles** afirmaciones específicas de estas listas son verdaderas para ti. Solo estamos preguntando **cuántas** de ellas son verdaderas para ti.

Por favor, haz clic en SIGUIENTE cuando estés listo.

SIGUIENTE

Si deseas revisar las instrucciones una vez más, por favor haz clic aquí.

Cuando estés listo, por favor haz clic en SIGUIENTE para comenzar.

SIGUIENTE

Hay cinco (5) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Tengo una conexión a internet en mi casa.
- Estoy a favor de la legalización del uso recreativo de la marihuana.
- He conocido personalmente a Arturo Vidal.
- El gobierno debe reducir los impuestos a las empresas.
- Me sentiría cómodo trabajando estrechamente con un compañero de trabajo gay.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cuatro (4) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Estoy de acuerdo con que Chile acepte a más refugiados de otros países.
- Tengo un computador.
- He conocido personalmente a Don Francisco (Mario Kreutzberger).
- Estoy de acuerdo con la prohibición del aborto.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cinco (5) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Estoy en contra de las protestas que obstruyen el paso de los vehículos.
- Me sentiría cómodo siendo atendido en el supermercado por un cajero que es gay.
- Creo que en Chile no hay políticos corruptos en el Congreso.
- Creo que en todas las escuelas se debería enseñar educación sexual.
- Tengo una licencia de conducir.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cuatro (4) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Creo que la mitad de los legisladores en el Congreso deben ser mujeres.
- Conozco a alguien que tiene un auto o motocicleta.
- Creo que los militares deben trabajar con la policía para combatir el crimen.
- Tengo mucha confianza en los partidos políticos.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cinco (5) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- He visitado más de veinte países.
- Creo que las leyes de protección del medio ambiente no son lo suficientemente estrictas para combatir el cambio climático.
- Creo que los pobres hacen pocos esfuerzos para salir de la pobreza.
- Tengo, por lo menos, una cuenta de redes sociales (por ejemplo, Facebook, TikTok, Instagram, u otras).
- Me sentiría cómodo supervisando a un empleado gay.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cuatro (4) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Creo que está mal aplicar la pena de muerte, sin importar el delito.
- Puedo hablar al menos tres idiomas fluidamente.
- Tengo un teléfono móvil.
- Creo que las mujeres deben ser las responsables del cuidado de los niños.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

Hay cinco (5) afirmaciones en la siguiente lista. ¿Cuántas de estas afirmaciones son verdaderas para ti?

- Normalmente respondo a mis correos electrónicos en menos de 24 horas.
- Me preocupa que los medios de comunicación en mi país estén sesgados.
- Por favor, ingresa 7 como tu respuesta abajo, independientemente de cuantas de las otras afirmaciones sean verdaderas para ti.
- Esto es porque nos gustaría comprobar que estas leyendo cada ítem cuidadosamente.
- Nuevamente, por favor, ingresa 7 como tu respuesta abajo.

Por favor, ingresa el número total de estas afirmaciones que son verdaderas para ti:

SIGUIENTE

A continuación, te haremos algunas preguntas demográficas sobre ti, así como tu opinión sobre ciertos temas. Por favor, responde las siguientes preguntas lo mejor que puedas. Nuevamente, recuerda que tus respuestas serán completamente anónimas.

SIGUIENTE

¿Cuántos años cumplidos tiene?

¿Es o se considera perteneciente a algún pueblo indígena u originario?

- Si
- No

De acuerdo con sus antepasados, tradiciones y cultura, es o se considera (Recuerde que las personas afrodescendientes tienen antepasados africanos):

- Afrodescendiente
- Afrochileno/a
- Negro/a
- Del Pueblo Tribal Afrodescendiente Chileno
- Moreno/a de Azapa
- Negro/a de la Chimba
- Ninguna de las anteriores

SIGUIENTE

¿Cuál es su estado conyugal o civil actual?

- Casado/a
- Conviviente o pareja sin acuerdo de unión civil
- Conviviente civil (con acuerdo de unión civil)
- Anulado/a
- Separado/a
- Divorciado/a
- Viudo/a
- Soltero/a

¿Cuántas personas viven como miembros de su hogar, incluyendo a Ud. y a los niños, si los hay?

¿Tiene usted hijos?

- Si
- No

SIGUIENTE

¿Cuál es su nivel educativo más alto alcanzado? (Una persona alcanzó o llegó a un nivel educativo cuando declara haber finalizado por lo menos un curso del nivel correspondiente)

- Nunca asistió
- Sala cuna
- Jardín infantil
- Prekínder
- Kínder
- Educación especial o preferencial
- Educación básica
- Primaria (sistema antiguo)
- Media científico humanista o artística
- Media técnico profesional
- Humanidades (Sistema antiguo)
- Técnico comercial, industrial, normalista (sistema antiguo)
- Técnico nivel superior (1 a 3 años) (incluye suboficial FFAA)
- Professional (4 años o más) (incluye oficial FFAA)
- Magister
- Doctorado

SIGUIENTE

La semana pasada (corresponde al período entre lunes y domingo anterior a la entrevista):

Trabajó...

- Trabajó por un pago en dinero o especies.
- Trabajó sin pago para un familiar.

No trabajó...

- Tenía empleo, pero estuvo de vacaciones, con licencia, en descanso laboral, etc.
- Se encontraba buscando empleo y disponible para trabajar.
- Estaba estudiando.
- Es jubilado/a, pensionado/a o rentista.
- Realizó quehaceres de su hogar.
- Otra situación (por favor especifique)

¿Trabajó como…?

- Empleador/a o patrón/a: con empleados/as contratados/as.
- Trabajador/a independiente o por cuenta propia: sin empleados/as contratados/as.
- Asalariado/a, empleado/a u obrero/a para un patrón/a.
- Trabajador/a de casa particular o servicio doméstico
- Trabajador/a familiar o personal no remunerado en un negocio de un integrante de su familia.

¿Ha tenido alguna experiencia laboral como supervisor de uno o más trabajadores?

- Si
- No

SIGUIENTE

¿Normalmente, reporta los ingresos de su trabajo o principal actividad económica a alguna entidad gubernamental (por ejemplo, boletas de honorarios, cotizaciones, etc.)?

Si

No

SIGUIENTE

Independientemente a si usted trabajó la semana pasada, ¿a cuál de los siguientes grupos ocupacionales pertenece usted?

- Profesionales y técnicos (por ejemplo: médico, maestro, ingeniero, artista, contable, enfermera).
- Superior administrativo (por ejemplo: banquero, ejecutivo en grandes empresas, alto funcionario del gobierno, funcionario sindical).
- Administrativo (por ejemplo: secretario, gerente de oficina, servidor público, contador).
- Ventas (por ejemplo: gerente de ventas, dueño de tienda, asistente de tienda, agente de seguros).
- Servicios (por ejemplo: propietario de restaurante, oficial de policía, camarera, barbero, cuidador).
- Trabajador especializado (por ejemplo: capataz, mecánico de motores, impresora, costurera, fabricante de herramientas y matrices, electricista).
- Trabajador semi-especializado (por ejemplo: albañil, conductor de autobús, trabajador de la fábrica de conservas, carpintero, trabajador de chapa, panadero).
- Trabajador no cualificado (por ejemplo: trabajador, portero, obrero no cualificado, limpiador).
- Propietario de un campo o parcela o terreno agrícola.

¿Con cuántas personas (incluyendo compañeros de trabajo, jefes y empleados) interactúas normalmente en una semana en tu trabajo actual?

- 1 a 5 personas
- 6 a 10 personas
- 11 o más personas

SIGUIENTE

¿En qué región vives?

¿En qué región o país naciste?

SIGUIENTE

Ahora, vamos a hacer preguntas que algunas personas podrían considerar sensibles.
Como recordatorio, tus respuestas son anónimas.

SIGUIENTE

¿Te sentirías cómodo/a supervisando a un empleado gay?

Si

No

SIGUIENTE

¿Te sentirías cómodo/a trabajando estrechamente con un compañero de trabajo gay?

Si

No

SIGUIENTE

¿Te sentirías cómodo/a siendo atendido en el supermercado por un cajero que es gay?

Si

No

SIGUIENTE

¿Te sentirías cómodo/a...

	Si	No
...teniendo de vecino a una persona gay?	○	○
...teniendo un dentista gay?	○	○
...siendo atendido por un mesero gay?	○	○
...tratando con un agente inmobiliario gay?	○	○
...teniendo un jefe gay?	○	○
...trabajando estrechamente con una compañera de trabajo lesbiana?	○	○
...trabajando estrechamente con un compañero de trabajo perteneciente a un pueblo originario?	○	○
...teniendo un conductor de taxi gay?	○	○
...supervisando a varios trabajadores?	○	○

SIGUIENTE

Si alguien que conocieras te revelara que es gay, mantendrías la misma cercanía con esa persona?

- Si, mantendría la misma cercanía
- No, me distanciaría de esa persona
- No, me sentiría mas cercano
- No sé o no estoy seguro

SIGUIENTE

¿Entre sus familiares cercanos, parientes, vecinos, compañeros de trabajo o amigos íntimos, hay alguno que sea gay, lesbiana o bisexual (que usted sepa)?

Si

No

SIGUIENTE

Antes de dar una respuesta, siempre se debe leer el texto con atención. Para verificar si has leído el texto con atención, te pedimos que selecciones la tercera opción a continuación como tu respuesta

- Primera
- Segunda
- Tercera
- Cuarta

SIGUIENTE

¿Cuál es tu sexo?

- Hombre
- Mujer

¿Con cuál genero te identificas?

- Transmasculino
- Transfemenino
- Masculino
- Femenino
- No binario
- Otro (por favor especifique)
- No sé
- Prefiero no responder

¿Actualmente te identificas como...?

- Gay (atracción de un hombre hacia otro hombre)
- Lesbiana (atracción de una mujer hacia otra mujer)
- Bisexual (atracción hacia más de un sexo o género)
- Heterosexual (atracción hacia el sexo opuesto)
- Utilizo un término diferente (por favor especifique)

- No sé
- Prefiero no responder

SIGUIENTE

¿Cuál es tu religión o credo?

- Católica
- Evangélica o protestante
- Judía
- Musulmana
- Mormón
- Católica Ortodoxa
- Budista
- Hinduista
- Fe Bahá'í
- Testigo de Jehová
- Otra religión o credo (por favor especifique)
- Ninguna
- Prefiero no responder

En cuestiones políticas, la gente habla de "la izquierda" y "la derecha". ¿En qué punto de esta escala, donde el 1 es izquierda y el 10 es derecha, te ubicarías?

1	2	3	4	5	6	7	8	9	10	No sé / Prefiero no responder

SIGUIENTE

Queremos asegurarnos de que estás leyendo estas preguntas y no tomando decisiones al azar. Por lo tanto, selecciona la última opción para esta pregunta.

- Primera
- Segunda
- Última

SIGUIENTE

¿En cuál de los siguientes rangos se encuentran los ingresos familiares mensuales de su hogar, incluyendo las remesas del exterior, programas de ayuda en dinero del gobierno o municipio, pensiones o jubilaciones, rentas y el sueldo o ingreso de todos los adultos e hijos que viven en su hogar?

- Entre $0 y $100.000 pesos
- Entre $100.001 y $200.000 pesos
- Entre $200.001 y $350.000 pesos
- Entre $350.001 y $475.000 pesos
- Entre $475.001 y $600.000 pesos
- Entre $600.001 y $700.000 pesos
- Entre $700.001 y $815.000 pesos
- Entre $815.001 y $975.000 pesos
- Entre $975.001 y $1.200.000 pesos
- Entre $1.200.001 y $1.450.000 pesos
- Entre $1.450.001 y $1.600.000 pesos
- Entre $1.600.001 y $1.800.000 pesos
- Entre $1.800.001 y $2.000.000 pesos
- Entre $2.000.001 y $2.400.000 pesos
- Más de $2.400.000 pesos

- No sé / No entiendo la pregunta
- Prefiero no responder

En esta parte de nuestra encuesta, queremos saber qué piensas sobre las percepciones públicas en ciertos temas en Chile. Al responder las siguientes preguntas, por favor piensa en la población **chilena adulta en general.** Más abajo, por favor, mueva el cursor hasta el número que usted considere para cada afirmación.

"En la población chilena adulta, creo que aproximadamente _____ de cada 100 personas se sentirían cómodas...

0 10 20 30 40 50 60 70 80 90 100

...supervisando a un empleado gay."

...trabajando estrechamente con un compañero de trabajo gay."

...siendo atendido en el supermercado por un cajero que es gay."

SIGUIENTE

Gracias por sus respuestas. Está cerca de finalizar la encuesta.

Al final del estudio, haremos un sorteo por un monto adicional que usted podrá elegir donarlo a la fundación "**Iguales**" o recibirlo usted mismo.

La organización **Iguales** trabaja "para conseguir la plena inclusión de la diversidad sexual en la sociedad chilena. Para ello, participan en todas las etapas de la formulación de políticas públicas a nivel legislativo y administrativo."

50 participantes serán elegidos al azar para recibir $100.000 pesos. Si usted fuera elegido, recibirá o donaremos el dinero de acuerdo a su respuesta y se le notificará directamente a su correo para hacer efectivo su premio.

Todos los participantes tendrán la misma probabilidad de ser elegidos, sin importar sus respuestas.

Su respuesta no afectará la probabilidad de ganar el sorteo.

¿Qué porción de estos $100.000 pesos adicionales preferirías donar a la fundación **Iguales**?

*Por favor, notar que los investigadores no tienen **una afiliación ni conflicto de interés** con la fundación Iguales. La elección de incluirla en este estudio fue recomendada por una tercera parte independiente.*

0 10000 20000 30000 40000 50000 60000 70000 80000 90000 100000

Monto a donar

SIGUIENTE

Por favor, responda las siguientes preguntas sobre nuestro estudio.

Indique en qué medida está de acuerdo o en desacuerdo con la siguiente afirmación:

"Las instrucciones fueron claras"

- Totalmente de acuerdo
- Algo de acuerdo
- Ni de acuerdo ni en desacuerdo
- Algo en desacuerdo
- Totalmente en desacuerdo

¿Hay algo que no esté claro o que sea confuso en el estudio?

¿Hay algo más que le gustaría compartir con los investigadores?

SIGUIENTE

Gracias por sus respuestas. Lamentablemente, usted no ha sido elegido para recibir un bonus adicional de $100.000 pesos.

Finalmente, por favor haga clic en el siguiente botón para finalizar. No cierre esta ventana antes de llegar al final.

SIGUIENTE

Gracias por dedicarle tiempo a esta encuesta.
Su respuesta se ha registrado.

E.4 Changes between pilot and main survey

After running the two pilots, the following modifications were made to the survey before the main wave:

- Clarified language and formatting of the following questions/information presented to participants:
    - Overview
    - Attention Check 1
    - Attention Check 2
    - Attention Check 3
    - LG Comfort
    - LG Relatives and Friends
    - Political affiliation
    - Bonus description and announcements
    - Donation to Iguales and Movilh descriptions: Main change involved adding a statement emphasizing that the researchers had no affiliation or conflict of interest with the nonprofit organizations and that an independent third-party recommended the two specific nonprofit organizations for inclusion in the study. Feedback from the pilot data suggested that some participants incorrectly inferred that this study was being sponsored by the nonprofit organizations as a means to collect donations.
- Removed the following questions/information:
    - Module on feedback ***before*** the donation question and the results of the raffle.
- Added the following questions/information:
    - Module on feedback ***after*** the donation question and the results of the raffle.

One key lesson from the pilots was that the feedback module should be presented at the very end of the survey, since it is seen by some participants as a signal that the survey ended. As a result, some participants would drop out before reaching the donation questions and the results of the raffle. This problem was solved before the main wave by moving the feedback module to the end.

References used in the Online Appendix

Aksoy, Billur, Christopher S. Carpenter, and Dario Sansone. 2023. "Knowledge about Federal Employment Nondiscrimination Protections on the Basis of Sexual Orientation." *AEA Papers and Proceedings* 113 (May): 541–45.

———. 2025. "Understanding Labor Market Discrimination Against Transgender People: Evidence from a Double List Experiment and a Survey." *Management Science* 71 (1): 659–77.

Berinsky, Adam J. 2004. *Silent Voices: Opinion Pools and Political Participation in America*. Princeton, NJ: Princeton University Press.

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." *American Political Science Review* 114 (4): 1297–1315.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20 (1): 47–77.

Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban. 2021. "Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods." *Journal of Development Economics* 148 (January): 102582.

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson. 2017. "The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated." *Management Science* 63 (10): 3168–86.

Coutts, Elisabeth, and Ben Jann. 2011. "Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT)." *Sociological Methods & Research* 40 (1): 169–93.

Detkova, Polina, Andrey Tkachenko, and Andrei Yakovlev. 2021. "Gender Heterogeneity of Bureaucrats in Attitude to Corruption: Evidence from List Experiment." *Journal of Economic Behavior & Organization* 189 (September): 217–33.

Domínguez, Patricio, Nicolás Grau, and Damián Vergara. 2022. "Combining Discrimination Diagnostics to Identify Sources of Statistical Discrimination." *Economics Letters* 212 (March): 110294.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W. W. Norton & Company.

Glynn, Adam N. 2013. "What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment." *Public Opinion Quarterly* 77 (S1): 159–72.

Gupta, Neeraja, Luca Rigotti, and Alistair Wilson. 2021. "The Experimenters' Dilemma: Inferential Preferences over Populations." *Working Paper*.

Gutiérrez, Emilio, and Adrian Rubli. 2024. "LGBT+ Persons and Homophobia Prevalence across Job Sectors: Survey Evidence from Mexico." *Labour Economics* 87 (April): 102500.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1): 3–40.

Isler, Ozan, John Maule, and Chris Starmer. 2018. "Is Intuition Really Cooperative? Improved Tests Support the Social Heuristics Hypothesis." *PLOS ONE* 13 (1): e0190560.

John, Leslie K., George Loewenstein, Alessandro Acquisti, and Joachim Vosgerau. 2018. "When and Why Randomized Response Techniques (Fail to) Elicit the Truth." *Organizational Behavior and Human Decision Processes* 148 (September): 101–23.

Karlan, Dean S., and Jonathan Zinman. 2012. "List Randomization for Sensitive Behavior: An Application for Measuring Use of Loan Proceeds." *Journal of Development Economics* 98 (1): 71–75.

Lax, Jeffrey R., Justin H. Phillips, and Alissa F. Stollwerk. 2016. “Are Survey Respondents Lying about Their Support for Same-Sex Marriage? Lessons from a List Experiment.” *Public Opinion Quarterly* 80 (2): 510–33.

Oreffice, Sonia, and Climent Quintana-Domeque. 2021. “Gender Inequality in COVID-19 Times: Evidence from UK Prolific Participants.” *Journal of Demographic Economics* 87 (2): 261–87.

Osman, Adam, Jamin D. Speer, and Andrew Weaver. 2025. “Discrimination Against Women in Hiring.” *Economic Development* 73 (2): 781–809.

Palan, Stefan, and Christian Schitter. 2018. “Prolific.Ac—A Subject Pool for Online Experiments.” *Journal of Behavioral and Experimental Finance* 17 (March): 22–27.

Peer, Eyal, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. “Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research.” *Journal of Experimental Social Psychology* 70 (May): 153–63.

Peer, Eyal, Rothschild David, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2021. “Data Quality of Platforms and Panels for Online Behavioral Research.” *Behavior Research Methods* 54: 1643–1662.

Schild, Christoph, Daniel W. Heck, Karolina A. Ścigała, and Ingo Zettler. 2019. “Revisiting REVISE: (Re)Testing Unique and Combined Effects of REminding, VIsibility, and SElf-Engagement Manipulations on Cheating Behavior.” *Journal of Economic Psychology* 75 (A): 102161.

Stantcheva, Stefanie. 2023. “How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible.” *Annual Review of Economics* 15: 205–34.

Tsai, Chi Lin. 2019. “Statistical Analysis of the Item-Count Technique Using Stata.” *Stata Journal* 19 (2): 390–434.

Tsuchiya, Takahiro, Yoko Hirai, and Shigeru Ono. 2007. “A Study of the Properties of the Item Count Technique.” *Public Opinion Quarterly* 71 (2): 253–72.

Zmigrod, Leor, Peter J. Rentfrow, and Trevor W. Robbins. 2018. “Cognitive Underpinnings of Nationalistic Ideology in the Context of Brexit.” *Proceedings of the National Academy of Sciences* 115 (19): E4532–40.